

Modeling extreme values with a GEV mixture probability distributions

Application to localisation w.r.t. longitude

Pascal Alain Dkengne Sielenou

2023-10-12

```
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)

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

## Rows: 20002 Columns: 25
## -- Column specification -----
## Delimiter: ","
## dbf (25): version_major, version_minor, status, timestamp, latitude, longitu...
##
## 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.
Gnss_map_matching <- xfun::in_dir(dir = path, expr = read_csv("./applications/Gnss_map_matching.csv"))

## Rows: 20001 Columns: 25
## -- Column specification -----
## Delimiter: ","
## dbf (25): version_major, version_minor, status, timestamp, latitude, longitu...
##
## 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.
longitude_Gnss_map_matching_errors <- Gnss_imar$longitude[-1] - Gnss_map_matching$longitude

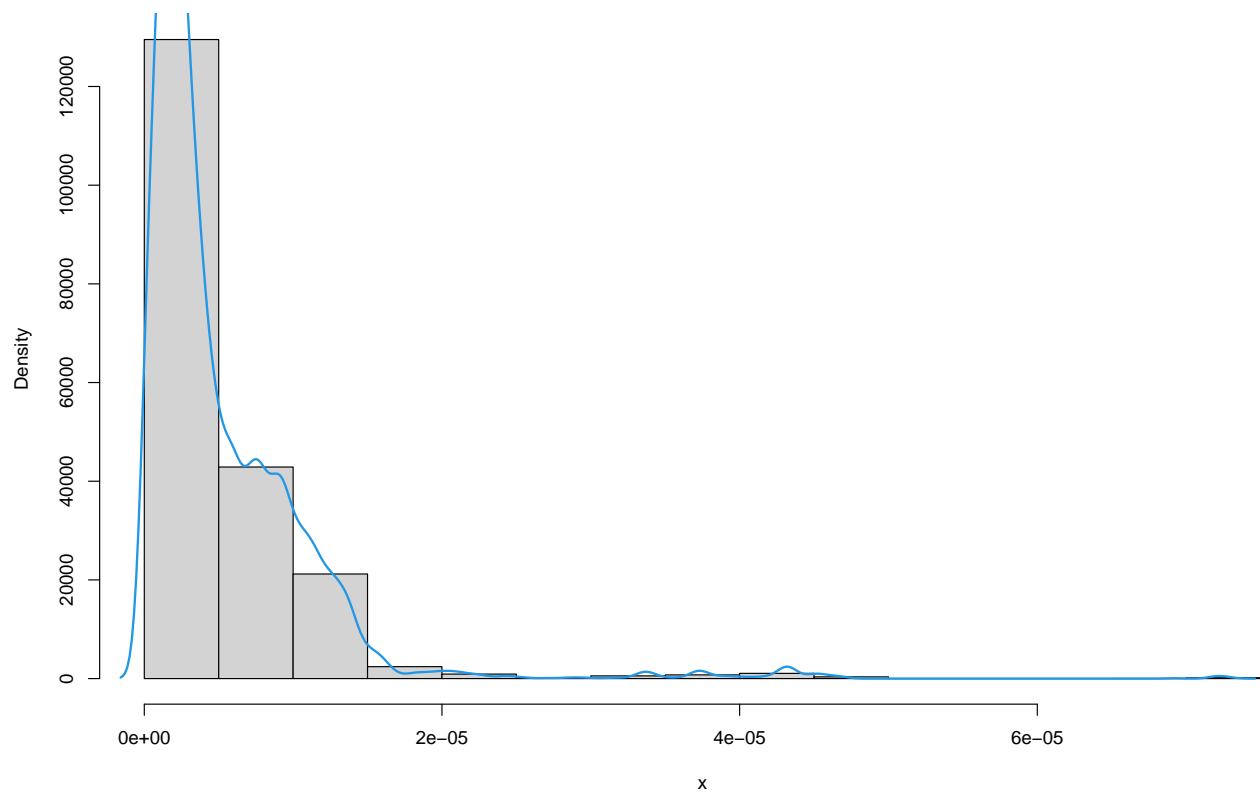
x <- abs(longitude_Gnss_map_matching_errors)
n <- length(x)

n

## [1] 20001
# Histogram of all data
```

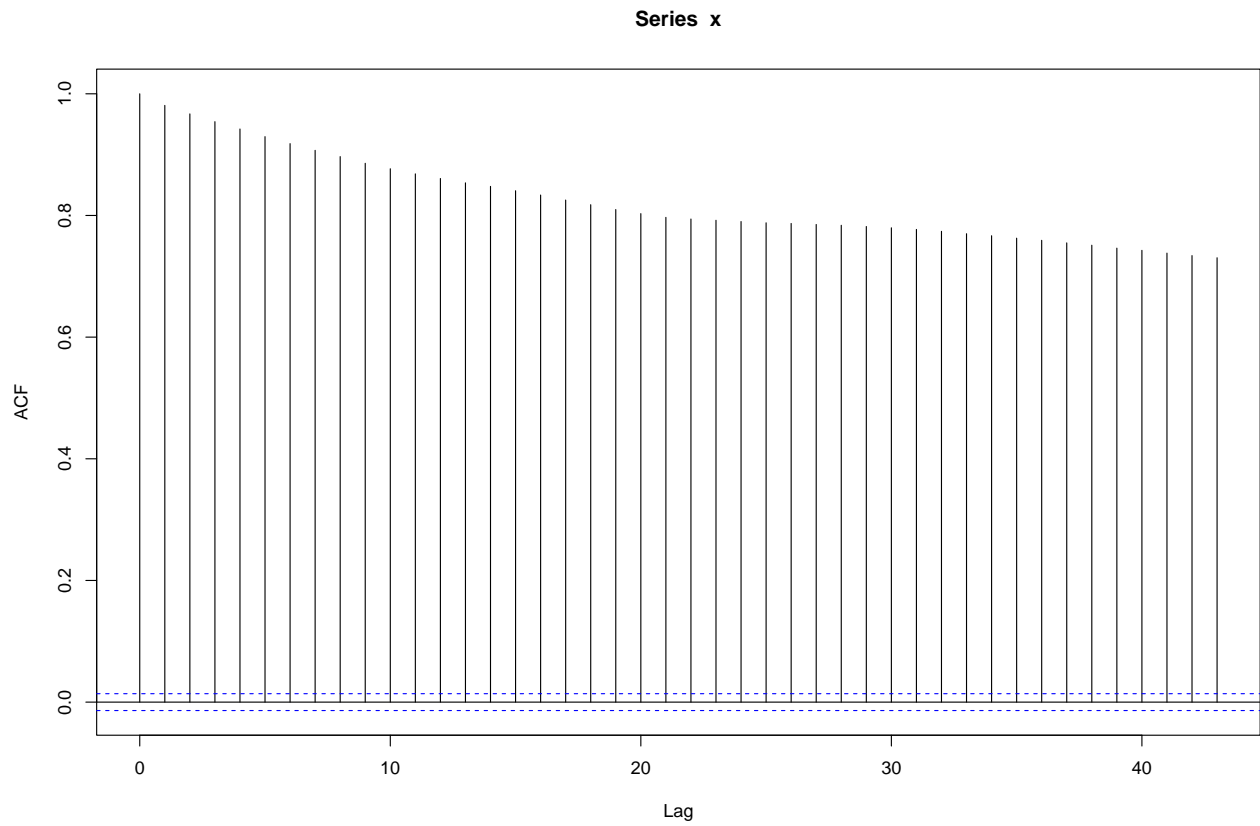
```
hist(x, prob = TRUE)
lines(density(x),
      lwd = 2,
      col = 4)
```

Histogram of x



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

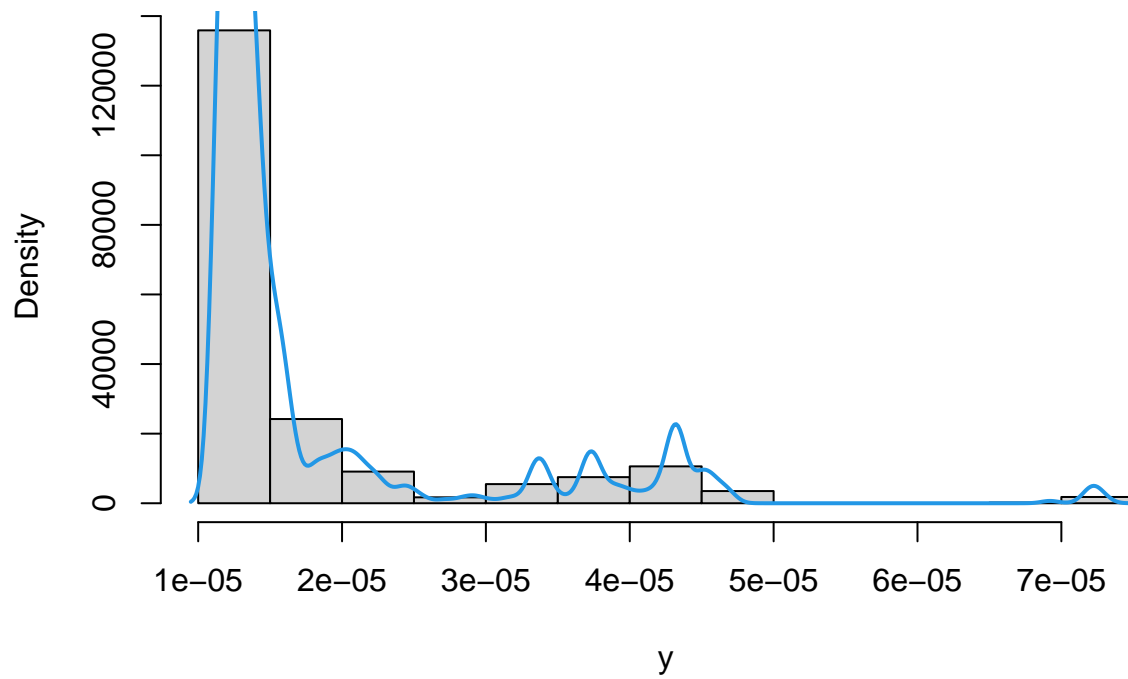
```
acf(x)
```



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

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

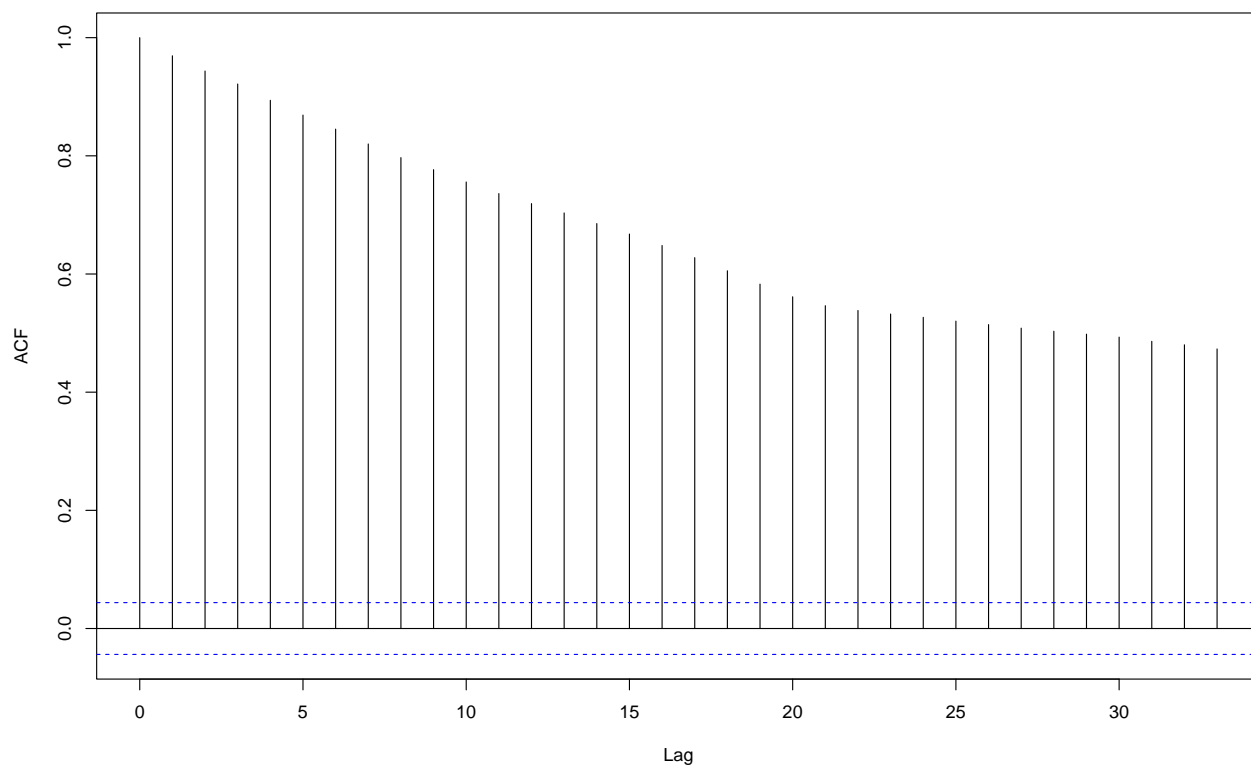
Histogram of y



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

```
acf(y)
```

Series y



```
y <- extract_nlargest_sample(x, n = nlargest)
```

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

```
gev_mixture_model <- suppressWarnings(estimate_gev_mixture_model_parameters(x = x,
                                                                              block_sizes = NULL,
                                                                              minimum_nblocks = 50,
                                                                              threshold = min(y),
                                                                              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
```

```
##           2           3           4           5           6
## 0.07431749242 0.15431147635 0.11256468399 0.10982777943 0.07708509135
##           7           8           9          10          11
## 0.10982777943 0.05340727771 0.05340727771 0.05340727771 0.03421679345
##          12          13          14          15          16
## 0.05340727771 0.05340727771 0.05340727771 0.04163206040 0.03313610154
##          17          18          19          20          21
## 0.03090271250 0.03421679345 0.03090271250 0.04163206040 0.03313610154
##          22          23          24          25          26
## 0.03165934408 0.03165934408 0.03165934408 0.03669113495 0.03165934408
##          27          28          29          30          31
## 0.03882452751 0.03669113495 0.03669113495 0.03165934408 0.02461729364
##          32          33          34          35          36
## 0.02614513731 0.03165934408 0.02614513731 0.02461729364 0.03165934408
##          37          38          39          40
## 0.02239998851 0.02277324991 0.02277324991 0.02277324991
```

```
gev_mixture_model$normalized_gev_parameters_object
```

```
##           loc_star      scale_star  shape_star
## 2  1.197774742e-05 1.030653382e-06 0.9392509589
## 3  1.172686857e-05 7.982592163e-07 0.9047005165
## 4  1.157525727e-05 6.591469518e-07 0.8898986007
## 5  1.147791624e-05 5.658144579e-07 0.8825950522
## 6  1.135716730e-05 5.004071990e-07 0.8817971075
## 7  1.131114547e-05 4.514647071e-07 0.8786317869
## 8  1.127055594e-05 4.056937682e-07 0.8953378625
## 9  1.121063000e-05 3.875689626e-07 0.8829650862
## 10 1.118425017e-05 3.689861448e-07 0.8775166949
## 11 1.115388837e-05 3.369452585e-07 0.8830703642
## 12 1.107849734e-05 3.408728285e-07 0.8709342811
## 13 1.108198444e-05 3.105473832e-07 0.8804293344
```

```
## 14 1.099750070e-05 3.100500806e-07 0.8769978235
## 15 1.106445810e-05 2.829672839e-07 0.8823880809
## 16 1.109383739e-05 2.433532173e-07 0.9151407604
## 17 1.090427577e-05 3.032639969e-07 0.8435206379
## 18 1.096477131e-05 2.475740949e-07 0.8882366631
## 19 1.085639316e-05 2.613145975e-07 0.8785151116
## 20 1.084028562e-05 2.843624375e-07 0.8383304760
## 21 1.095672588e-05 2.208401509e-07 0.9056342085
## 22 1.087282048e-05 2.423235289e-07 0.8701367064
## 23 1.089410916e-05 2.223323535e-07 0.8868842511
## 24 1.093645869e-05 1.925758760e-07 0.9359241819
## 25 1.081687565e-05 2.259039728e-07 0.8667340229
## 26 1.088434586e-05 1.972530773e-07 0.9028741883
## 27 1.093823762e-05 1.764990106e-07 0.9344045661
## 28 1.073588973e-05 1.991065582e-07 0.9157044866
## 29 1.068275782e-05 2.169616806e-07 0.8941103808
## 30 1.068444812e-05 2.239283264e-07 0.8753635363
## 31 1.088574829e-05 1.671275669e-07 0.9262894276
## 32 1.098154317e-05 1.553572762e-07 0.9252778334
## 33 1.076954153e-05 1.831614213e-07 0.8817846758
## 34 1.039575561e-05 2.691369479e-07 0.8002701405
## 35 1.066680764e-05 1.755110424e-07 0.9216080879
## 36 1.048790160e-05 2.160446714e-07 0.8612322331
## 37 1.069669898e-05 1.640864827e-07 0.9254615793
## 38 1.095741713e-05 1.338209625e-07 0.9689244251
## 39 1.099368552e-05 1.488659170e-07 0.9211229694
## 40 1.077857612e-05 1.547977205e-07 0.9261009816
```

```
gev_mixture_model$full_normalized_gev_parameters_object
```

```
##          loc_star      scale_star  shape_star
## 2  1.097593299e-05 8.969821829e-08 0.9392509589
## 3  1.100722033e-05 1.471930815e-07 0.9047005165
## 4  1.094060227e-05 9.436834845e-08 0.8898986007
## 5  1.092808934e-05 8.053995397e-08 0.8825950522
## 6  1.084890441e-05 5.222245202e-08 0.8817971075
## 7  1.087110148e-05 6.482806427e-08 0.8786317869
## 8  1.085032181e-05 2.944223912e-08 0.8953378625
## 9  1.080472065e-05 2.916517621e-08 0.8829650862
## 10 1.079591275e-05 2.821358045e-08 0.8775166949
## 11 1.079170020e-05 1.710761710e-08 0.8830703642
## 12 1.071761898e-05 2.657149498e-08 0.8709342811
## 13 1.075600264e-05 2.354344318e-08 0.8804293344
## 14 1.067103829e-05 2.374325218e-08 0.8769978235
## 15 1.076317794e-05 1.712126561e-08 0.8823880809
## 16 1.083968417e-05 1.076724516e-08 0.9151407604
## 17 1.056389688e-05 1.614737104e-08 0.8435206379
## 18 1.069995296e-05 1.235272617e-08 0.8882366631
## 19 1.057296626e-05 1.231977757e-08 0.8785151116
## 20 1.052469394e-05 1.979231374e-08 0.8383304760
## 21 1.072401897e-05 1.009281647e-08 0.9056342085
## 22 1.060813655e-05 1.201233010e-08 0.8701367064
## 23 1.065514873e-05 1.040210964e-08 0.8868842511
## 24 1.073882583e-05 7.606508620e-09 0.9359241819
## 25 1.057109321e-05 1.287597652e-08 0.8667340229
```

```
## 26 1.067554597e-05 8.733044418e-09 0.9028741883
## 27 1.075842365e-05 8.480015512e-09 0.9344045661
## 28 1.052899560e-05 9.652676823e-09 0.9157044866
## 29 1.045273564e-05 1.129645706e-08 0.8941103808
## 30 1.044109048e-05 1.090192239e-08 0.8753635363
## 31 1.071115747e-05 5.405937181e-09 0.9262894276
## 32 1.081940356e-05 5.333084980e-09 0.9252778334
## 33 1.057171573e-05 8.721668394e-09 0.8817846758
## 34 1.007765430e-05 1.456996598e-08 0.8002701405
## 35 1.048263540e-05 5.776416505e-09 0.9216080879
## 36 1.024986981e-05 1.104402734e-08 0.8612322331
## 37 1.052466814e-05 4.878547356e-09 0.9254615793
## 38 1.082284178e-05 3.427623590e-09 0.9689244251
## 39 1.083703180e-05 4.568585178e-09 0.9211229694
## 40 1.061646022e-05 4.662021290e-09 0.9261009816
```

```
gev_mixture_model$automatic_weights_pw_shape
```

```
##          2          3          4          5          6
## 3.330669074e-15 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##          7          8          9         10         11
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         12         13         14         15         16
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         17         18         19         20         21
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         22         23         24         25         26
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         27         28         29         30         31
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         32         33         34         35         36
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         37         38         39         40
## 0.000000000e+00 1.000000000e+00 0.000000000e+00 0.000000000e+00
```

```
gev_mixture_model$automatic_weights_pw_scale
```

```
##          2          3          4          5
## 0.000000000e+00 1.000000000e+00 -9.992007222e-16 0.000000000e+00
##          6          7          8          9
## 0.000000000e+00 0.000000000e+00 -2.775557562e-17 0.000000000e+00
##         10         11         12         13
## -2.775557562e-17 6.938893904e-17 -2.775557562e-17 0.000000000e+00
##         14         15         16         17
## 0.000000000e+00 -1.665334537e-16 -4.857225733e-17 -5.551115123e-17
##         18         19         20         21
## -4.163336342e-17 2.775557562e-17 -2.775557562e-17 -8.326672685e-17
##         22         23         24         25
## 0.000000000e+00 -7.632783294e-17 2.775557562e-17 8.326672685e-17
##         26         27         28         29
## -1.526556659e-16 -9.714451465e-17 2.775557562e-17 1.387778781e-16
##         30         31         32         33
## 4.857225733e-17 4.163336342e-17 -1.353084311e-16 -6.938893904e-17
##         34         35         36         37
## 2.775557562e-17 4.163336342e-17 1.110223025e-16 1.387778781e-17
```

```

##          38          39          40
## 1.110223025e-16 6.938893904e-17 -3.469446952e-18
gev_mixture_model$automatic_weights_pw_loc

##          2          3          4          5          6
## 5.329070518e-15 1.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##          7          8          9         10         11
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         12         13         14         15         16
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         17         18         19         20         21
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         22         23         24         25         26
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         27         28         29         30         31
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         32         33         34         35         36
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         37         38         39         40
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
gev_mixture_model$weighted_normalized_gev_parameters_object[3, ]

##          loc_star      scale_star  shape_star
## automatic_weights 1.172686857e-05 7.982592163e-07 0.9689244251
gev_mixture_model$automatic_weights_mw

##          2          3          4          5
## 0.000000000e+00 1.000000000e+00 0.000000000e+00 0.000000000e+00
##          6          7          8          9
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##         10         11         12         13
## 0.000000000e+00 0.000000000e+00 1.776356839e-15 1.776356839e-15
##         14         15         16         17
## 0.000000000e+00 -3.552713679e-15 1.776356839e-15 0.000000000e+00
##         18         19         20         21
## -8.881784197e-16 -1.776356839e-15 -1.776356839e-15 -8.881784197e-16
##         22         23         24         25
## 0.000000000e+00 0.000000000e+00 3.552713679e-15 -2.664535259e-15
##         26         27         28         29
## 0.000000000e+00 -2.664535259e-15 0.000000000e+00 0.000000000e+00
##         30         31         32         33
## 0.000000000e+00 1.332267630e-15 4.440892099e-16 8.881784197e-16
##         34         35         36         37
## 8.881784197e-16 -8.881784197e-16 3.552713679e-15 8.881784197e-16
##         38         39         40
## 8.881784197e-16 0.000000000e+00 2.220446049e-15

# Model diagnostics

## GEV mixture model with respect to parameters

par(mfrow = c(2, 1))

plot_gev_mixture_model_pdf(gev_mixture_model,

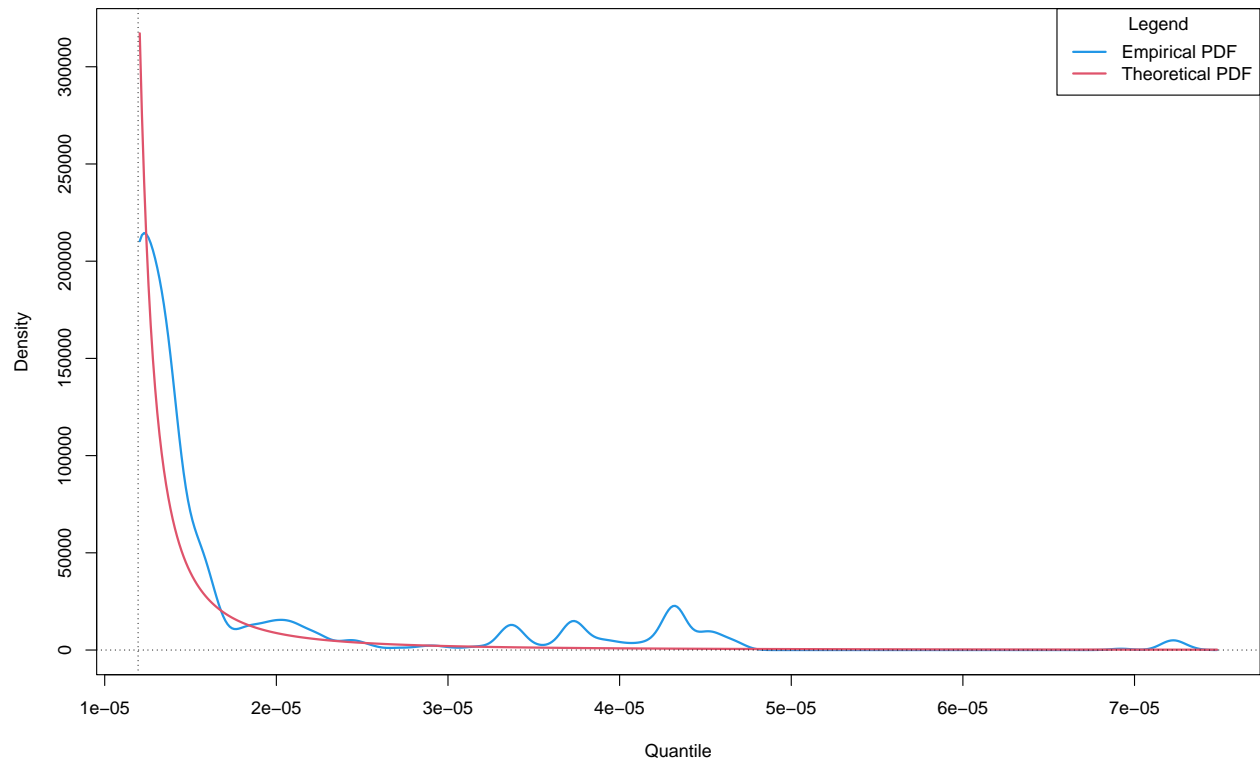
```



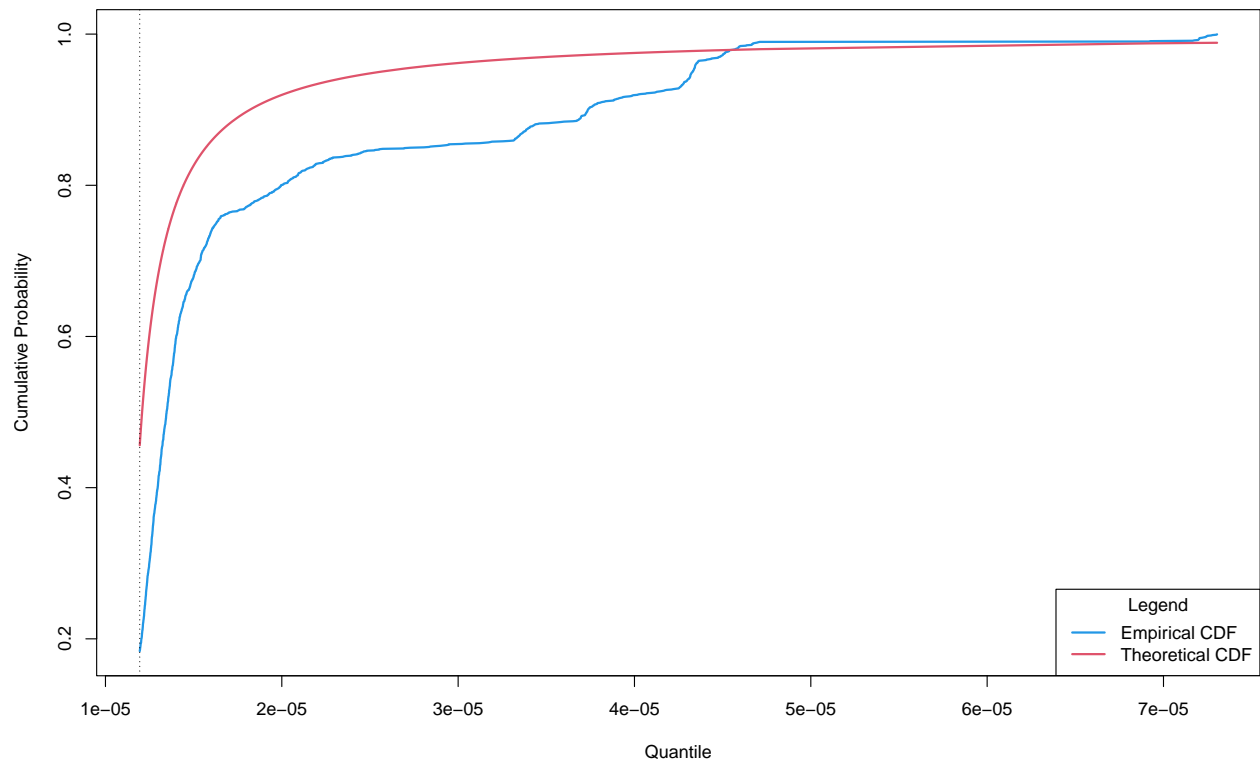
```
        type = "automatic_weights",
        model_wise = FALSE,
        zoom = TRUE,
        xlab = "Quantile",
        ylab = "Density",
        main = "Probability Density Function (PDF) Plot")

plot_gev_mixture_model_cdf(gev_mixture_model,
        type = "automatic_weights",
        model_wise = FALSE,
        zoom = TRUE,
        xlab = "Quantile",
        ylab = "Cumulative Probability",
        main = "Cumulative Distribution Function (CDF) Plot")
```

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



Cumulative Distribution Function (CDF) Plot : automatic_weights – model_wise = FALSE : zoom = TRUE

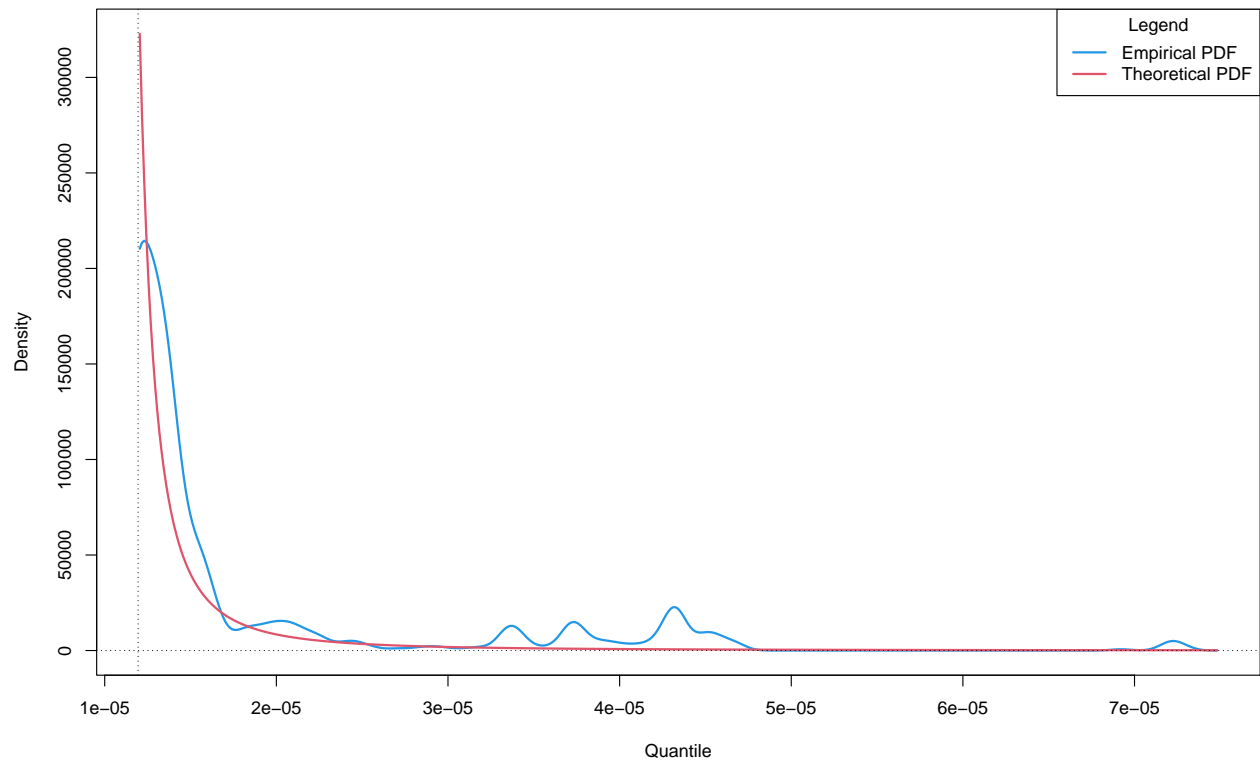


```
## GEV mixture model with respect to distribution functions
```

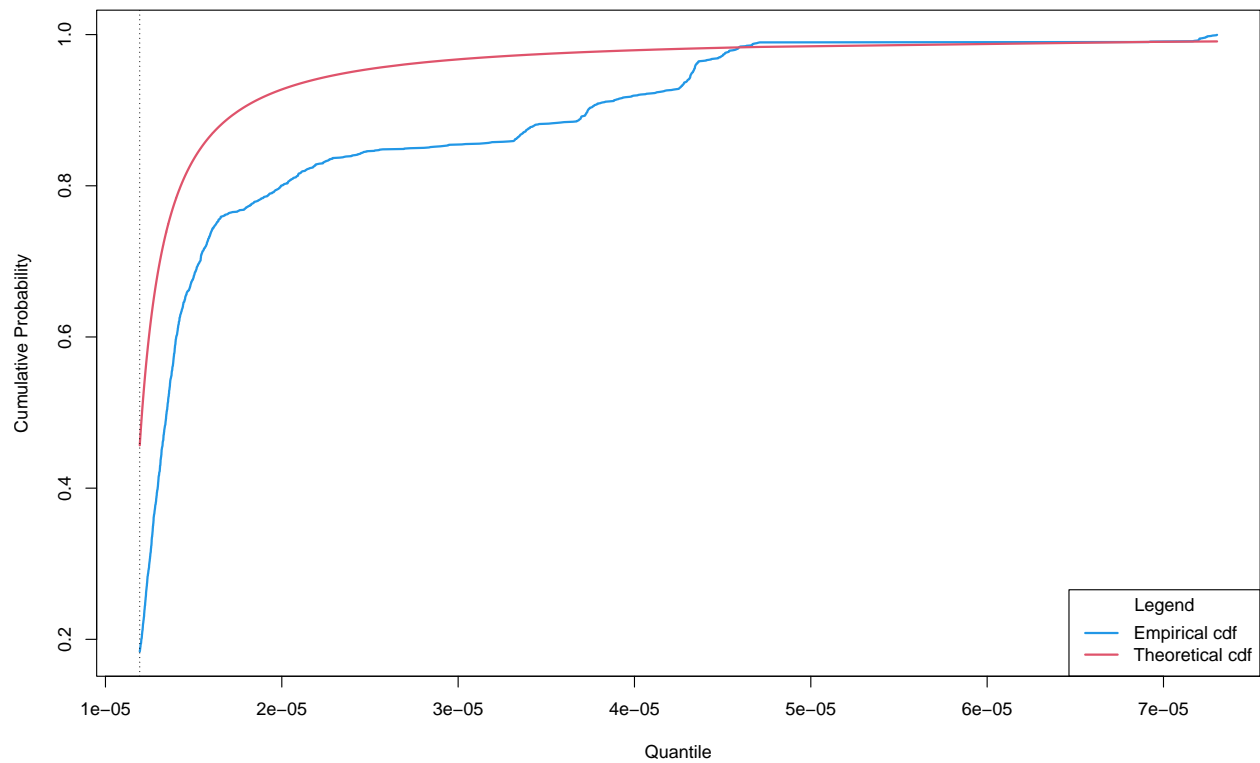
```
par(mfrow = c(2, 1))
plot_gev_mixture_model_pdf(gev_mixture_model,
                             type = "automatic_weights",
                             model_wise = TRUE,
                             zoom = TRUE,
                             xlab = "Quantile",
                             ylab = "Density",
                             main = "Probability Density Function (PDF) Plot")

plot_gev_mixture_model_cdf(gev_mixture_model,
                             type = "automatic_weights",
                             model_wise = TRUE,
                             zoom = TRUE,
                             xlab = "Quantile",
                             ylab = "Cumulative Probability",
                             main = "Cumulative Distribution Function (CDF) Plot")
```

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



Cumulative Distribution Function (CDF) 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",  
  "empirical",  
  "confidence_interval_mw",  
  "confidence_interval_pw")
```

```
alpha <- 10(-14)
```

```
## 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[2]
```

```
##      estimate  
## 1 3248932.877
```

```
## 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[2]
```

```
##      estimate  
## 1 508634.3495
```

```
## Quantiles from equivalent estimated GEV models
```

```
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[9]))
```

```
est_rl_pw
```

```
##      lower      estimate      upper  
## 2 -755836.9471 1780191.129 4316219.204  
## 3 -343418.0368 509060.0911 1361538.219  
## 4 -249603.0566 274353.5782 798310.2129  
## 5   -204162.4 190815.7227 585793.8454  
## 6 -212480.5487 164918.3808 542317.3102  
## 7 -188529.9251 135841.514 460212.9531
```

```
## 8 -314538.7077 197510.6946 709560.0968
## 9 -231288.8803 132107.3407 495503.5617
## 10 -201668.8737 107507.2682 416683.4101
## 11 -233773.333 115206.1453 464185.6236
## 12 -174126.4116 82175.05904 338476.5297
## 13 -214967.9565 98399.70279 411767.362
## 14 -210482.8871 88997.3798 388477.6467
## 15 -244332.4891 94867.1689 434066.8269
## 16 -539857.768 209684.2489 959226.2658
## 17 -92380.79748 33225.78282 158832.3631
## 18 -273297.6671 98226.42759 469750.5223
## 19 -229265.4518 78357.65445 385980.7607
## 20 -87033.2076 26837.43497 140708.0775
## 21 -439750.3595 144661.0446 729072.4487
## 22 -199502.2001 57090.40527 313683.0106
## 23 -296895.2762 84840.68818 466576.6526
## 24 -1076456.231 302247.0054 1680950.242
## 25 -176122.6373 48257.36164 272637.3606
## 26 -428811.6149 119327.2232 667466.0613
## 27 -1011322.711 265118.8569 1541560.424
## 28 -711686.4875 174371.4065 1060429.3
## 29 -424186.9611 101955.3848 628097.7308
## 30 -259360.1769 61323.4848 382007.1465
## 31 -817710.3677 198627.6507 1214965.669
## 32 -793737.0884 179332.027 1152401.142
## 33 -269858.6264 60347.07665 390552.7797
## 34 -43294.8111 8515.913838 60326.63877
## 35 -887890.8513 182238.623 1252368.097
## 36 -188744.0205 39392.9306 267529.8817
## 37 -965206.8434 190413.8389 1346034.521
## 38 -2935131.969 544793.0703 4024718.11
## 39 -745178.3317 152424.3564 1050027.044
## 40 -1058363.567 182976.1078 1424315.783
```

```
## Comparison of estimated quantiles
```

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

```
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[8]))
```

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

```
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)),
        cex = 1,
        cex.lab = 1,
        cex.axis = 1,
```

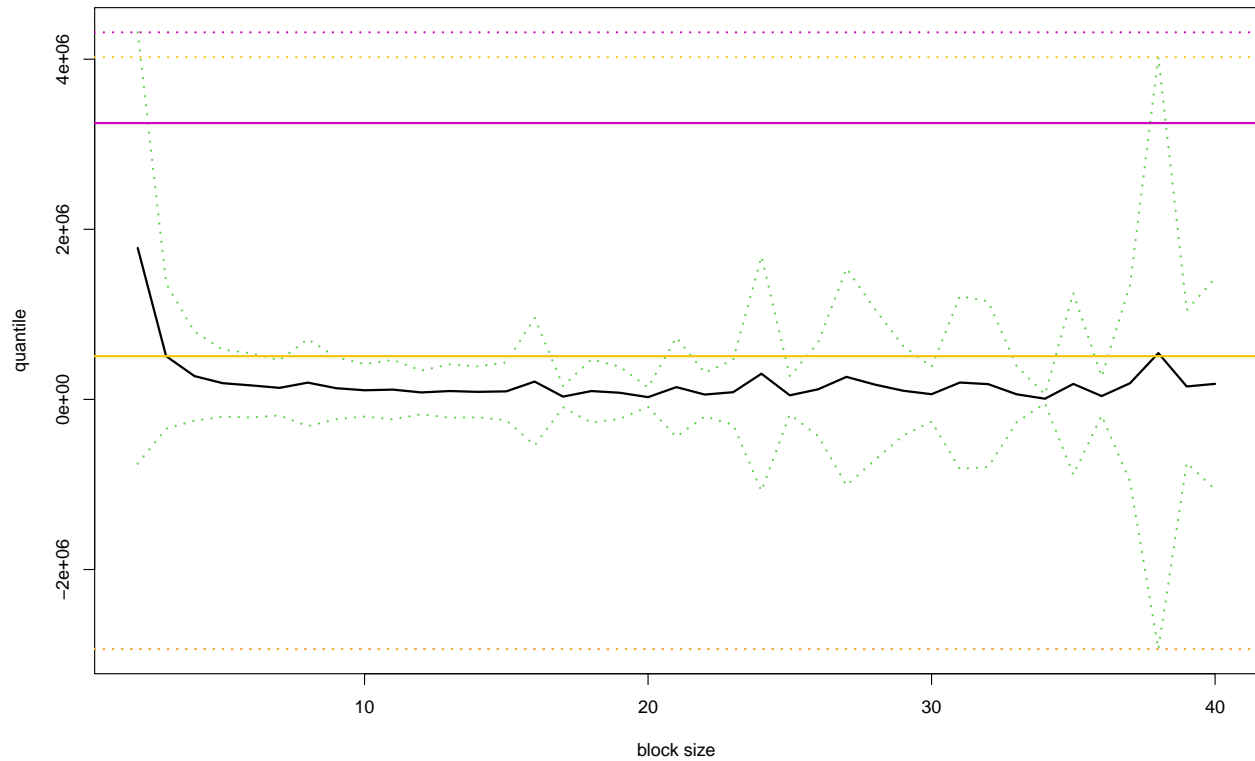
```

type = "l",
lty = c("dotted", "solid", "dotted"),
lwd = c(2,2,2),
col = c(3, 1, 3))

abline(h = rl_mw[2], col = 7, lwd = 2)
abline(h = rl_pw[2], 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)

```

Estimates of a quantile



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