

Modeling extreme values with a GEV mixture probability distributions

Application to localisation w.r.t. longitude

Pascal Alain Dkengne Sielenou

2023-10-26

```
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: ","
## dbl (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: ","
## dbl (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.
timestamp_position <- sapply(Gnss_map_matching$timestamp,
                             function(ts)
                               which.min(abs(ts - Gnss_imar$timestamp)))

longitude_Gnss_map_matching_errors <- Gnss_imar$longitude[-1] - Gnss_map_matching$longitude

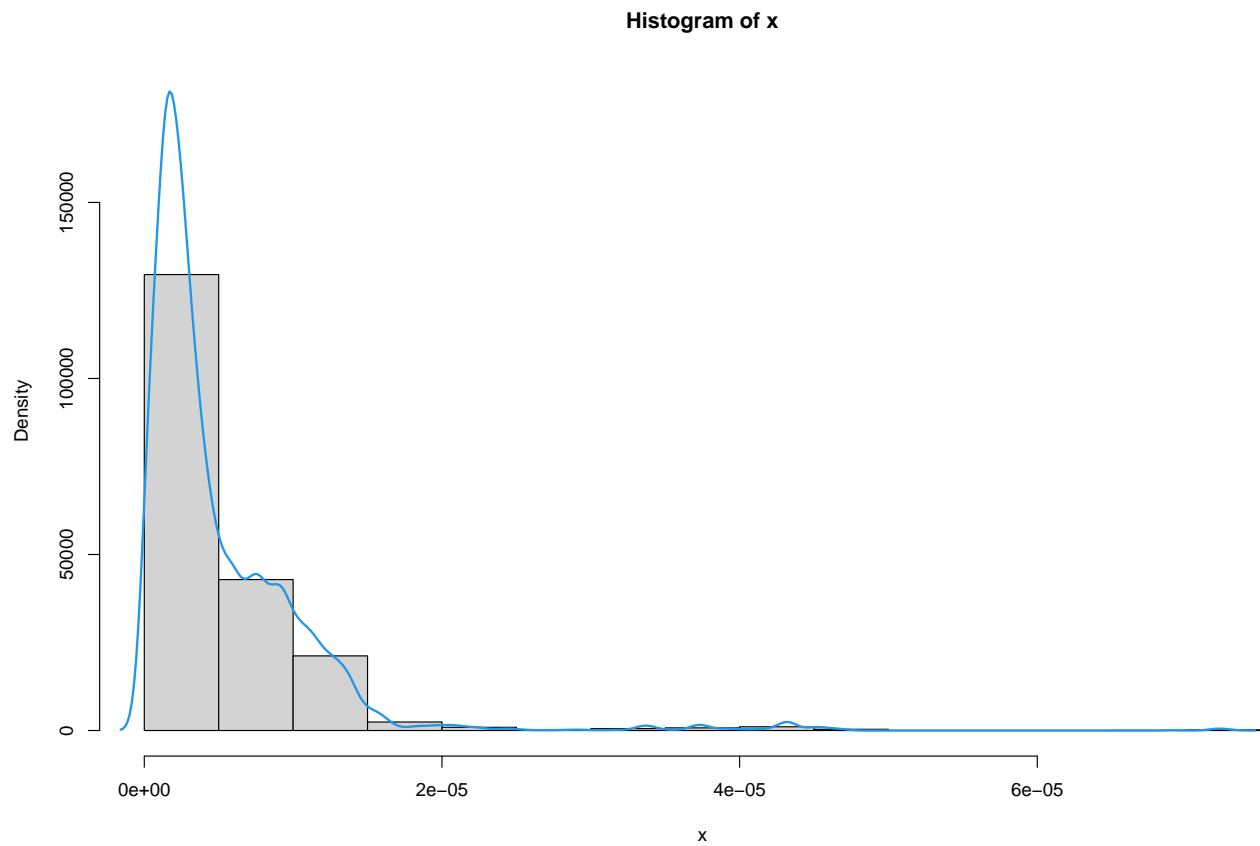
x <- abs(longitude_Gnss_map_matching_errors)
n <- length(x)

n
```

```
## [1] 20001
```

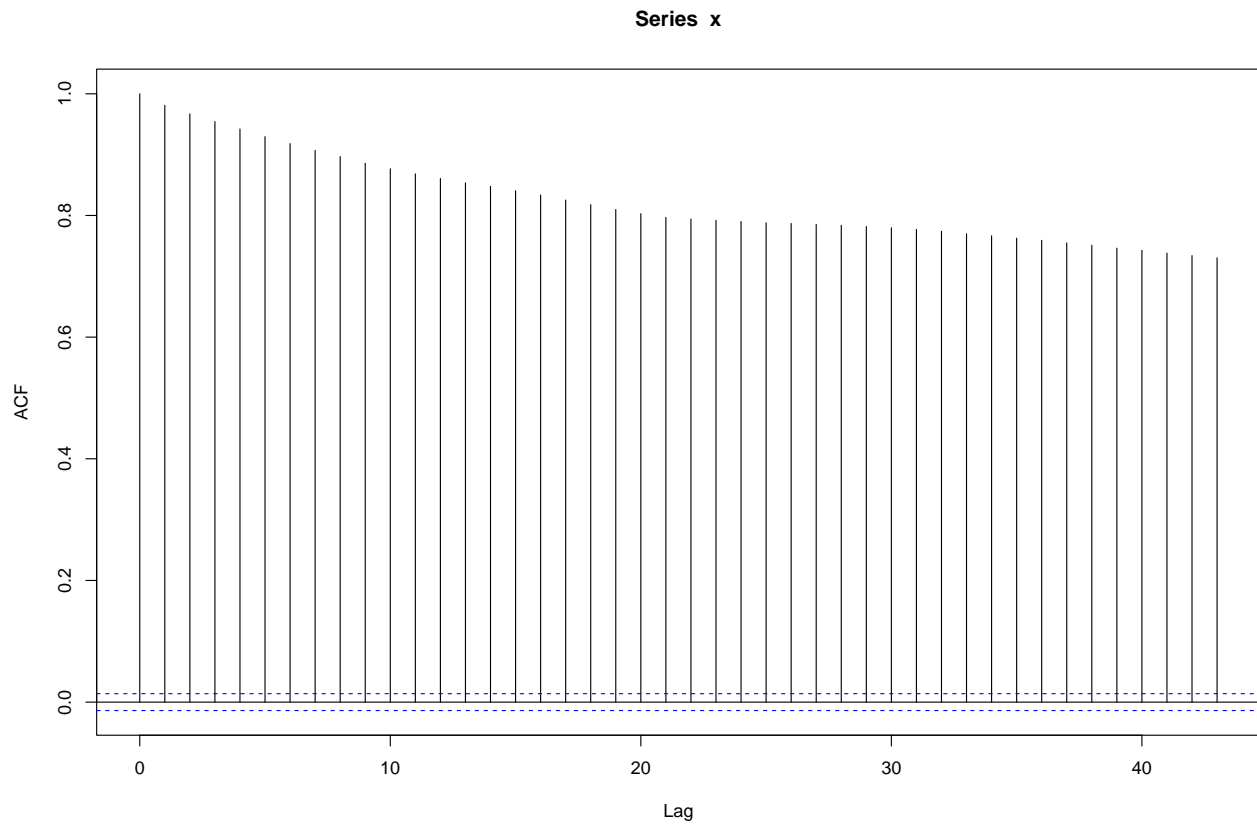
```
# Histogram of all data
```

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



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

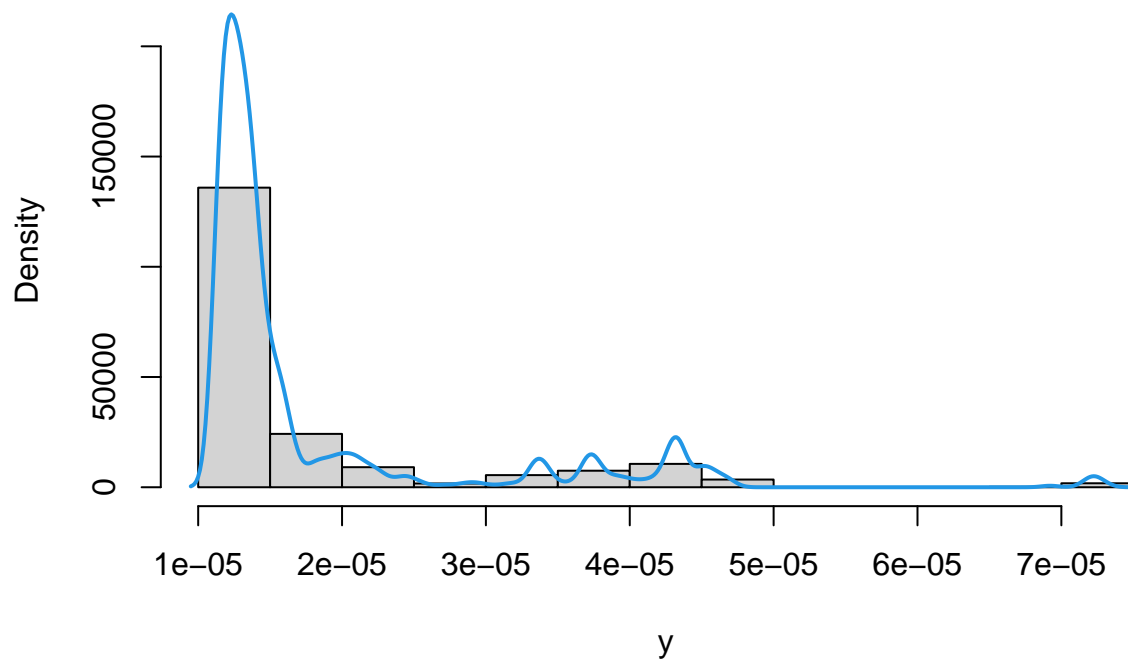
```
acf(x)
```



```
# 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)
```

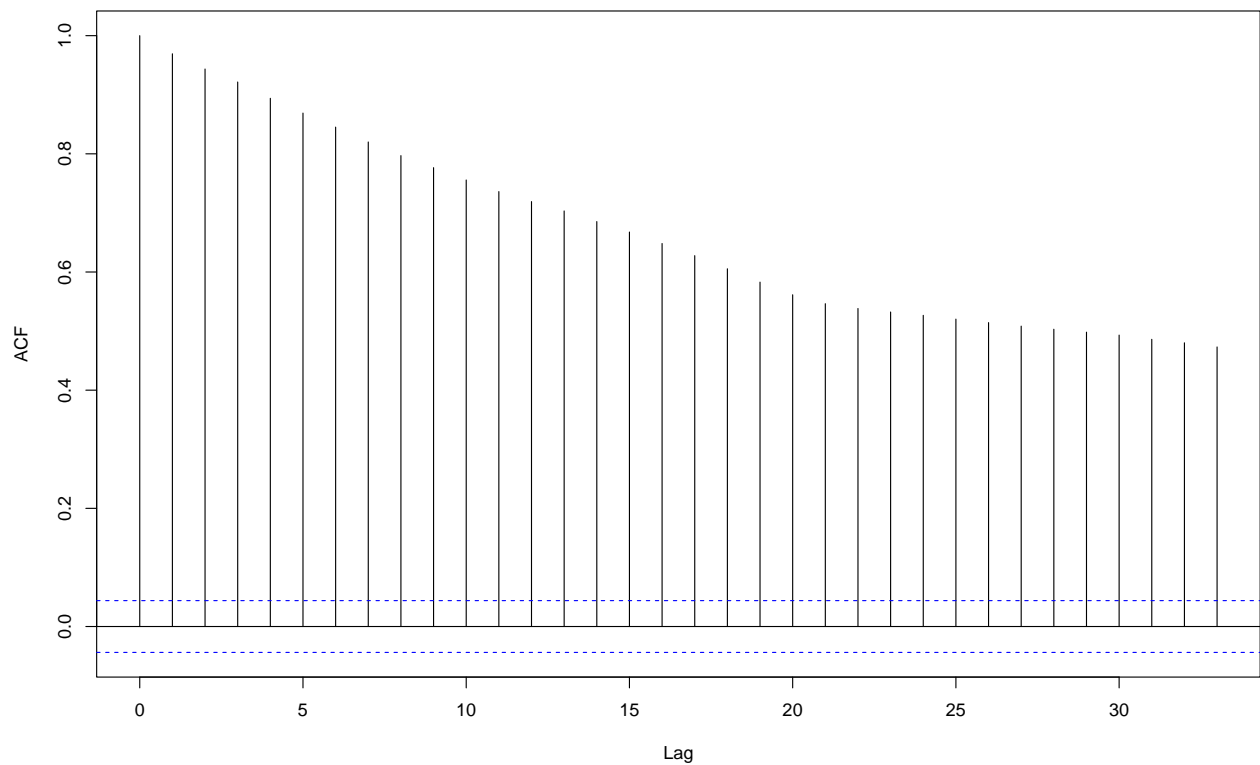
Histogram of y



```
# 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 = 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
```

```
##           10           11           12           13           14  
## 0.05340727771 0.03421679345 0.05340727771 0.05340727771 0.05340727771  
##           15           16           17           18           19  
## 0.04163206040 0.03313610154 0.03090271250 0.03421679345 0.03090271250  
##           20           21           22           23           24  
## 0.04163206040 0.03313610154 0.03165934408 0.03165934408 0.03165934408  
##           25           26           27           28           29  
## 0.03669113495 0.03165934408 0.03882452751 0.03669113495 0.03669113495  
##           30           31           32           33           34  
## 0.03165934408 0.02461729364 0.02614513731 0.03165934408 0.02614513731  
##           35           36           37           38           39  
## 0.02461729364 0.03165934408 0.02239998851 0.02277324991 0.02277324991  
##           40  
## 0.02277324991
```

```
gev_mixture_model$normalized_gev_parameters_object
```

```
##           loc_star      scale_star  shape_star  
## 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
## 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
```

```
## 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 0 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 1 0 0
```

```
gev_mixture_model$automatic_weights_pw_scale
```

```
##          10          11          12          13
## 1.000000000e+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 -8.881784197e-16 0.000000000e+00 8.881784197e-16
##          22          23          24          25
## 0.000000000e+00 8.881784197e-16 1.776356839e-15 0.000000000e+00
##          26          27          28          29
## -1.332267630e-15 1.332267630e-15 1.776356839e-15 -8.881784197e-16
##          30          31          32          33
## -1.776356839e-15 1.332267630e-15 4.440892099e-16 4.440892099e-16
##          34          35          36          37
## 0.000000000e+00 0.000000000e+00 -8.881784197e-16 -8.881784197e-16
##          38          39          40
## -8.881784197e-16 1.998401444e-15 -2.220446049e-16
```

```
gev_mixture_model$automatic_weights_pw_loc
```

```
##          10          11          12          13          14
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##          15          16          17          18          19
## 0.000000000e+00 1.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00
##          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 0.000000000e+00 0.000000000e+00
##          35          36          37          38          39
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 8.526512829e-14
##          40
## 0.000000000e+00
```

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

```
##          loc_star      scale_star  shape_star
## automatic_weights 1.109383739e-05 3.689861448e-07 0.9689244251
```

```
gev_mixture_model$automatic_weights_mw
```

```
##          10          11          12          13
## 1.000000000e+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 -8.881784197e-16 0.000000000e+00 0.000000000e+00
##          22          23          24          25
## 8.881784197e-16 0.000000000e+00 8.881784197e-16 -1.776356839e-15
##          26          27          28          29
## 8.881784197e-16 -8.881784197e-16 0.000000000e+00 8.881784197e-16
##          30          31          32          33
## 1.776356839e-15 8.881784197e-16 8.881784197e-16 1.776356839e-15
```

```

##          34          35          36          37
## 8.881784197e-16 8.881784197e-16 0.000000000e+00 -1.332267630e-15
##          38          39          40
## -8.881784197e-16 4.440892099e-16 -8.881784197e-16

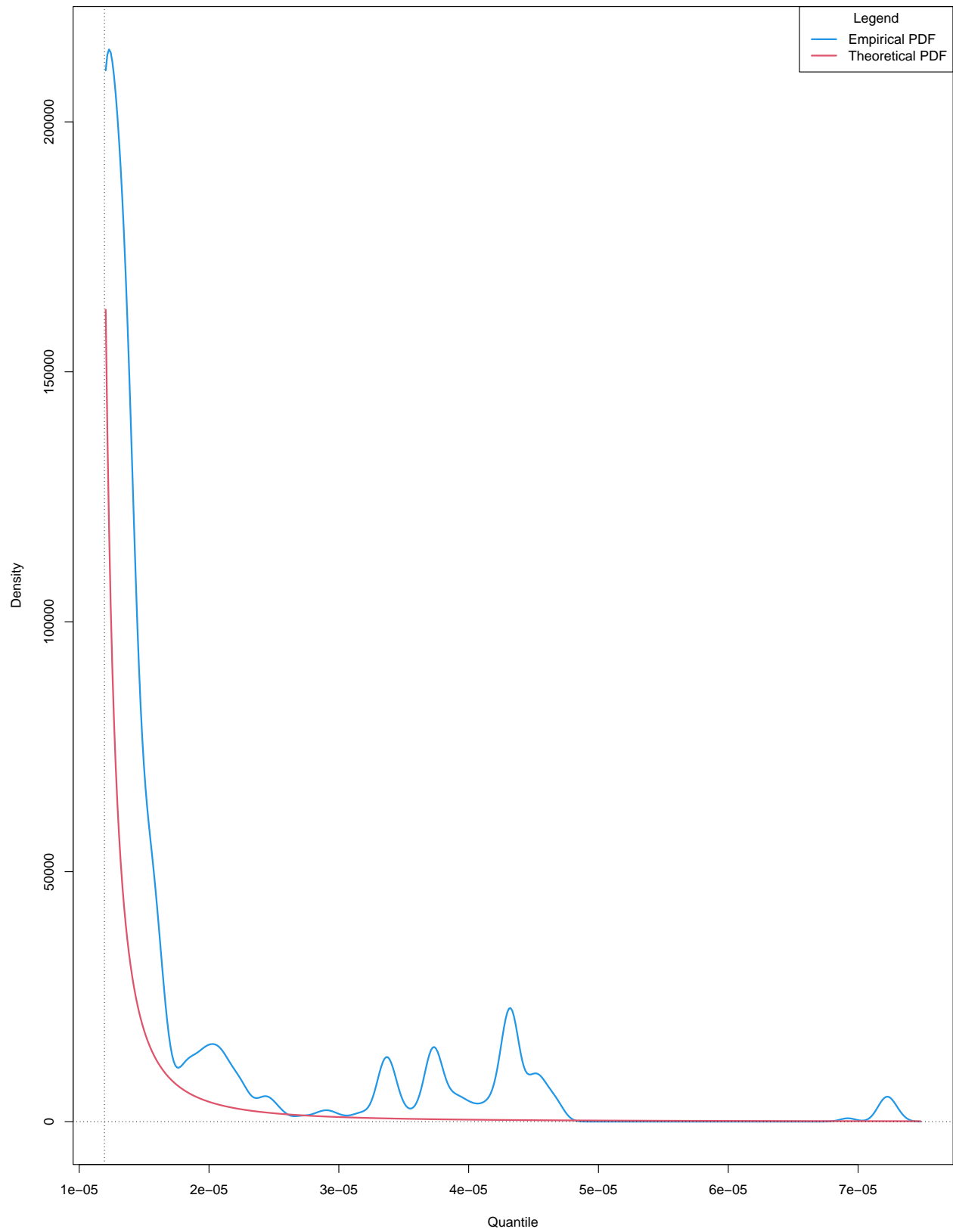
# Model diagnostics

## 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")

```

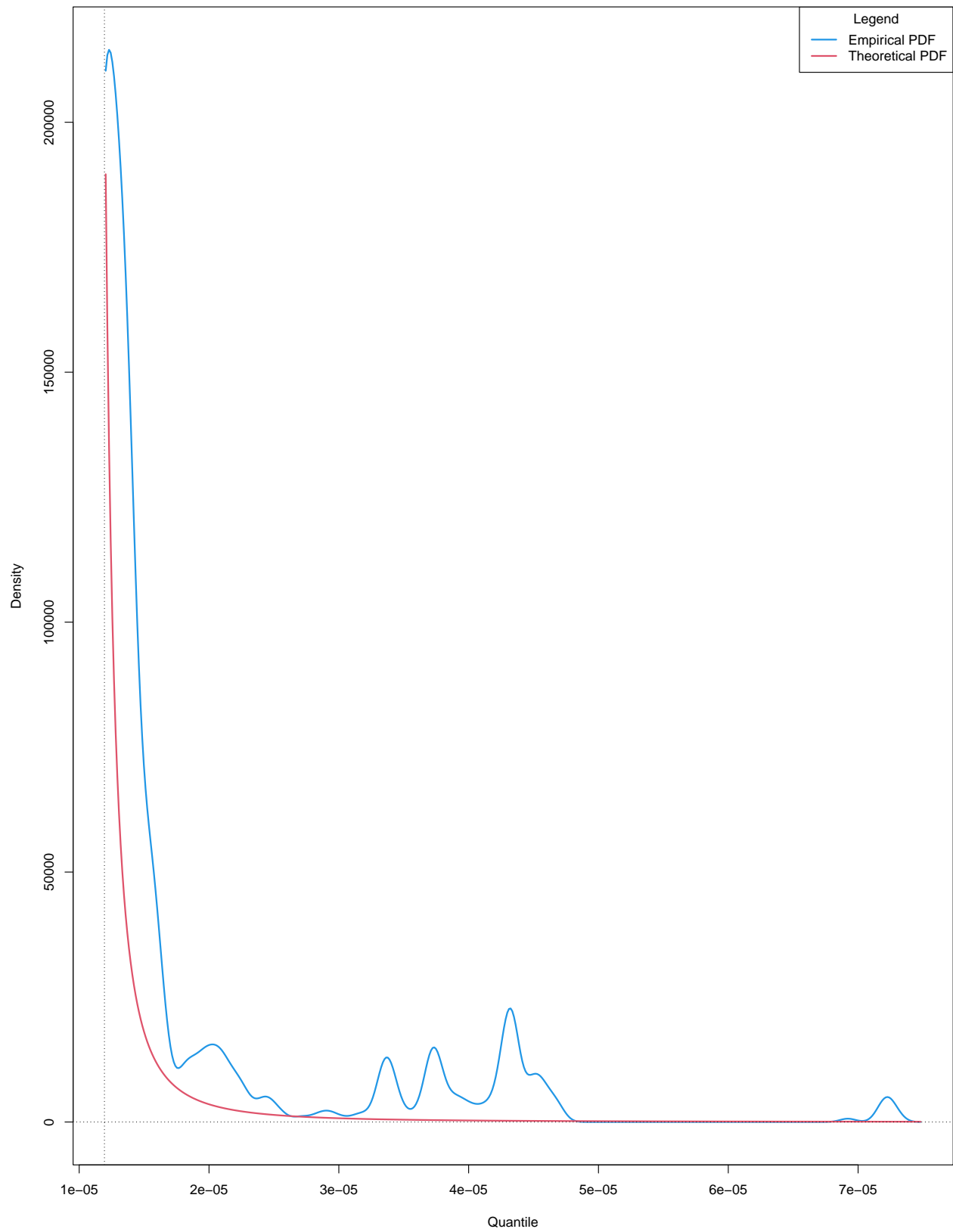

Probability Density Function (PDF) Plot : automatic_weights – model_wise = FALSE : zoom = TRUE



```
## 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^(-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
```

```
## [1] 42115.00888
```

```
## 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] 322.8917029
```

```
## 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 -681.9841868 330.92063420 1343.8254552  
## 11 -613.5909124 228.87655568 1071.3440237  
## 12 -598.1011877 257.66351163 1113.4282109  
## 13 -835.7117132 338.81867284 1513.3490589  
## 14 -704.6126289 268.17017907 1240.9529870  
## 15 -624.3544511 198.53535297 1021.4251570  
## 16 -1574.4426128 394.15784842 2362.7583096  
## 17 -267.5791979  79.02368590  425.6265697
```

```
## 18 -710.3242961 184.57067713 1079.4656503
## 19 -520.4267570 130.81781374 782.0623845
## 20 -303.4618963 84.78554447 473.0329852
## 21 -1063.5679697 234.78047938 1533.1289285
## 22 -510.6025128 114.23109555 739.0647039
## 23 -425.1796621 87.62842844 600.4365190
## 24 -2399.4719298 413.05688716 3225.5857041
## 25 -504.7280431 112.50013974 729.7283225
## 26 -1014.5294384 189.02745905 1392.5843564
## 27 -6998.1439838 1286.60932628 9571.3626364
## 28 -1514.7676169 276.44280662 2067.6532302
## 29 -1013.0236514 196.14173386 1405.3071191
## 30 -608.6204050 114.86372539 838.3478557
## 31 -1783.0889941 227.48643096 2238.0618561
## 32 -4797.9391446 602.79299855 6003.5251418
## 33 -663.4397696 109.63009709 882.6999638
## 34 -140.1860860 24.97269843 190.1314829
## 35 -1587.4964446 199.56542781 1986.6273002
## 36 -391.9808068 67.38587001 526.7525468
## 37 -1343.5866252 147.77326627 1639.1331577
## 38 -4458.0787954 368.25197978 5194.5827549
## 39 -1817.5595662 194.18313378 2205.9258337
## 40 -1786.6058506 182.26709817 2151.1400470
```

```
## 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
## 10 -681.9841868 330.92063420 1343.8254552
## 22 -510.6025128 114.23109555 739.0647039
## 24 -2399.4719298 413.05688716 3225.5857041
## 26 -1014.5294384 189.02745905 1392.5843564
## 29 -1013.0236514 196.14173386 1405.3071191
## 30 -608.6204050 114.86372539 838.3478557
## 31 -1783.0889941 227.48643096 2238.0618561
## 32 -4797.9391446 602.79299855 6003.5251418
## 33 -663.4397696 109.63009709 882.6999638
## 34 -140.1860860 24.97269843 190.1314829
## 35 -1587.4964446 199.56542781 1986.6273002
## 39 -1817.5595662 194.18313378 2205.9258337
```

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

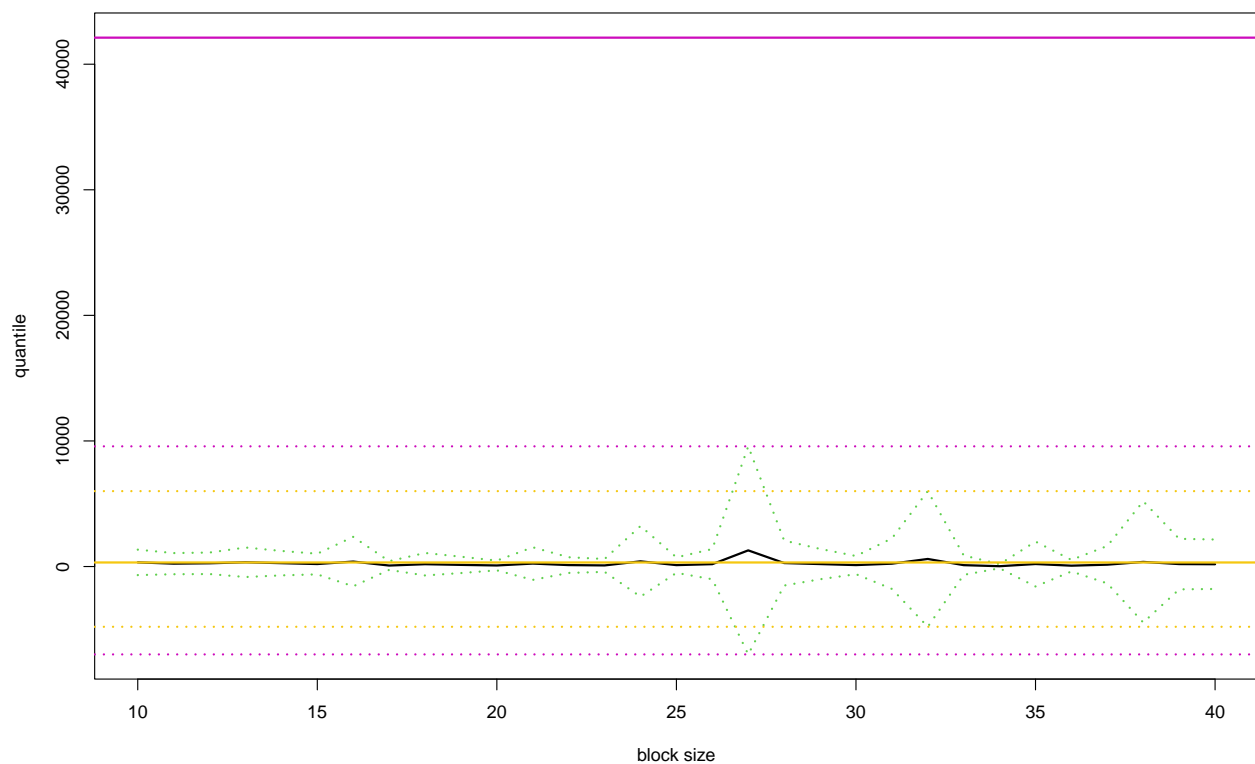
```
est_rl_mw_range
```

```
## [1] -4797.939145 6003.525142
```

```
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)
```

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
```