Nach Sonnenschein kommt Regen

Wetterdaten mit `rdwd` downloaden und verarbeiten, seltene Hochwässer mit `extremeStat` schätzen

Berry Boessenkool, uni-potsdam.de, March 2017

berry-b@gmx.de

github.com/brry/rdwd
cran.r-project.org/package=extremeStat

Nach Sonnenschein kommt Regen

Wetterdaten mit `rdwd` downloaden und verarbeiten, seltene Hochwässer mit `extremeStat` schätzen

Berry Boessenkool, uni-potsdam.de, March 2017

berry-b@gmx.de

github.com/brry/rdwd
cran.r-project.org/package=extremeStat

The German Weather Service (DWD) provides over 25'000 climate datasets

The German Weather Service (DWD) provides over 25'000 climate datasets

- Too much for manual inspection
- Somewhat difficult to search
- File format inconsistent (e.g. column widths)

The German Weather Service (DWD) provides over 25'000 climate datasets

- Too much for manual inspection
- Somewhat difficult to search
- File format inconsistent (e.g. column widths)

Screenshot of FTP server:

Index von ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany/climate/daily/more_precip/recent/

In den übergeordneten Ordner wechseln

Name	Größe	Zuletzt verändert
BESCHREIBUNG_obsgermany_climate_daily_more_precip_recent_de.pdf	67 KB	25.04.2016 00:00:00
DESCRIPTION_obsgermany_climate_daily_more_precip_recent_en.pdf	66 KB	25.04.2016 00:00:00
RR_Tageswerte_Beschreibung_Stationen.txt	1094 KB	27.01.2017 09:45:00
tageswerte_RR_00015_akt.zip	5 KB	27.01.2017 05:12:00
tageswerte_RR_00019_akt.zip	6 KB	26.01.2017 00:13:00
tageswerte_RR_00020_akt.zip	6 KB	26.01.2017 00:13:00
tageswerte_RR_00021_akt.zip	6 KB	26.01.2017 00:13:00
tageswerte_RR_00022_akt.zip	6 KB	26.01.2017 00:13:00
tageswerte_RR_00023_akt.zip	6 KB	26.01.2017 13:17:00
tageswerte_RR_00041_akt.zip	6 KB	26.01.2017 13:17:00
tageswerte_RR_00044_akt.zip	6 KB	26.01.2017 13:17:00
tageswerte_RR_00053_akt.zip	6 KB	27.01.2017 06:33:00

R saves the day

R package `rdwd` $->$ easy usage of the datasets

Select and download data

Select and download data

```
library("rdwd")
link <- selectDWD("Potsdam", res="daily",
                  var="kl", per="recent")
```

Select and download data

```
library("rdwd")
link <- selectDWD("Potsdam", res="daily",
                  var="kl", per="recent")
```

```
## ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany/
## /climate/daily/kl/recent/tageswerte_KL_03987_akt.zip
```

Select and download data

```
library("rdwd")
link <- selectDWD("Potsdam", res="daily",
                  var="kl", per="recent")
```

```
## ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany/
## /climate/daily/kl/recent/tageswerte_KL_03987_akt.zip
```

```
clim <- dataDWD(link)
```

```
## dataDWD -> dirDWD: creating directory 'C:/Users/boessenkool/rdwd/DWDdata'
## dataDWD -> fileDWD: creating 1 file: 'daily_kl_recent_tageswerte_KL_03987_akt.zip'
```

Select and download data

```
library("rdwd")
link <- selectDWD("Potsdam", res="daily",
                  var="kl", per="recent")
```

```
## ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany/
## /climate/daily/kl/recent/tageswerte_KL_03987_akt.zip
```

```
clim <- dataDWD(link)
```

```
## dataDWD -> dirDWD: creating directory 'C:/Users/boessenkool/rdwd/DWDdata'
## dataDWD -> fileDWD: creating 1 file: 'daily_kl_recent_tageswerte_KL_03987_akt.zip'
```

`dataDWD` uses `readDWD` to unzip file(s) + read / convert data

Data structure

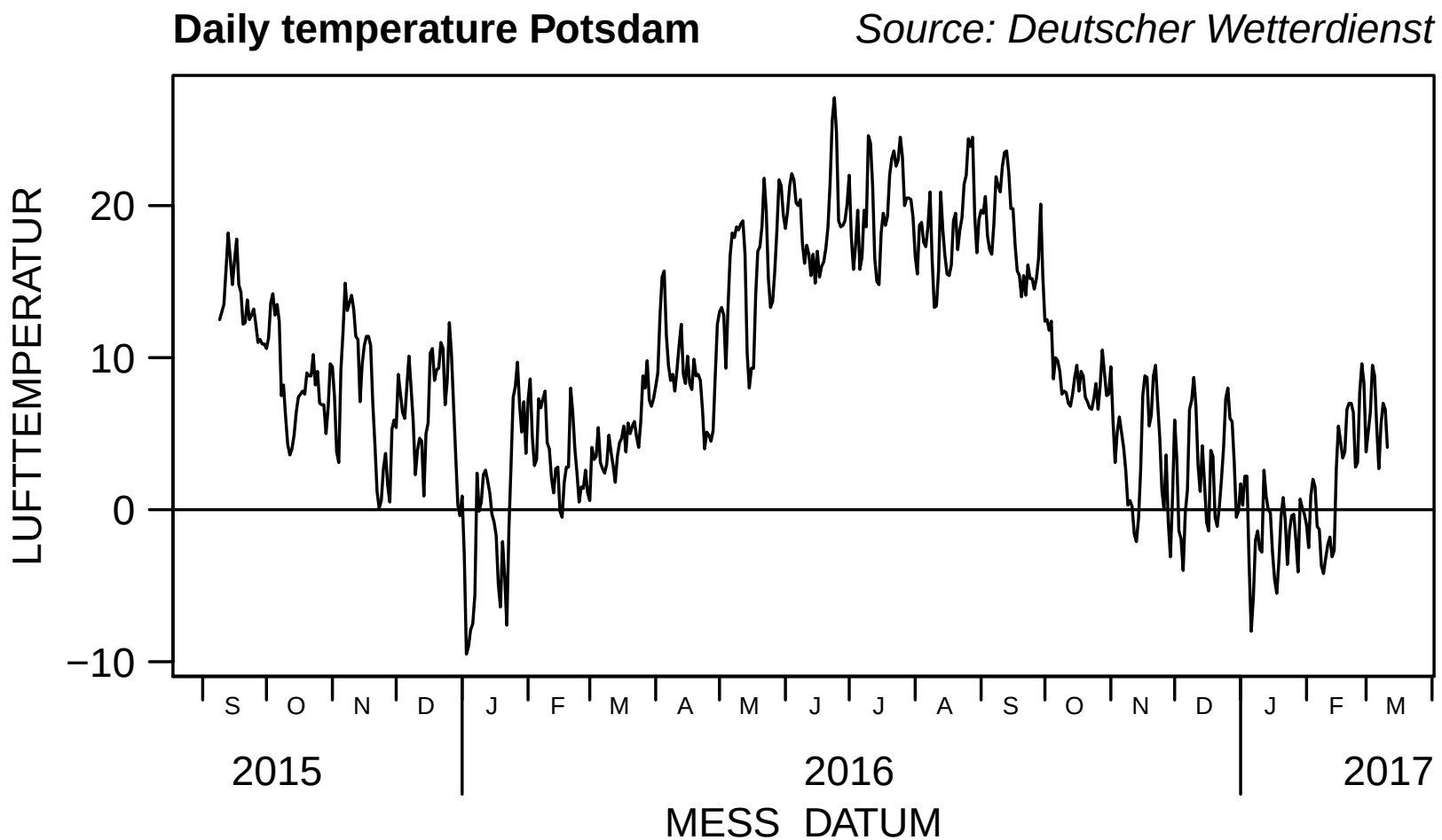
```
str(clim)
```

```
## 'data.frame': 550 obs. of  18 variables:
##  $ STATIONS_ID             : int  3987 3987 3987 3987 3987 3987 3987 3987 3987 3987 ...
##  $ MESS_DATUM              : POSIXct, format: "2015-09-09" "2015-09-10" ...
##  $ QUALITAETS_NIVEAU       : int  3 3 3 3 3 3 3 3 3 3 ...
##  $ LUFTTEMPERATUR          : num  12.5 13 13.5 15.8 18.2 16.5 14.8 16.4 17.8 14.8 ...
##  $ DAMPFDRUCK              : num  11.1 11.7 14.3 14.7 15.9 16.5 13.2 14.7 16.5 13.2 ...
##  $ BEDECKUNGSGRAD          : num  4.1 5 6.2 4.5 6.7 5.3 3.8 6 5.8 3.5 ...
##  $ LUFTDRUCK_STATIONSHOEHE : num  1013 1013 1009 1006 1000 ...
##  $ REL_FEUCHTE             : num  77.7 81 92.2 83.8 76.8 ...
##  $ WINDGESCHWINDIGKEIT     : num  3 3.6 4.1 3.4 3.8 4.4 4.5 4.2 5.1 4.4 ...
##  $ LUFTTEMPERATUR_MAXIMUM  : num  18.7 19.9 17.3 22.7 24.1 20.2 19.4 23.7 24.1 19.9 ...
##  $ LUFTTEMPERATUR_MINIMUM  : num  7.1 8.1 10.5 11.8 14.1 12.4 10.6 10.8 12.9 10.1 ...
##  $ LUFTTEMP_AM_ERDB_MINIMUM: num  4.5 5.2 8.8 8.7 13.5 10.9 8.4 8.1 11.5 7.9 ...
##  $ WINDSPITZE_MAXIMUM      : num  9.2 19.5 10.4 9.5 8.4 11.7 14.9 12.2 12.7 11.9 ...
##  $ NIEDERSCHLAGSHOEHE      : num  0 1.3 0.4 0.1 0 4.7 0 1.7 2.9 0 ...
##  $ NIEDERSCHLAGSHOEHE_IND  : int  6 6 6 6 0 6 6 6 6 0 ...
##  $ SONNENSCHEINDAUER       : num  6.82 5.7 1.55 6.12 5.32 ...
##  $ SCHNEEHOEHE             : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ eor                     : Factor w/ 1 level "eor": 1 1 1 1 1 1 1 1 1 1 ...
```

Data can be plotted with regular R code

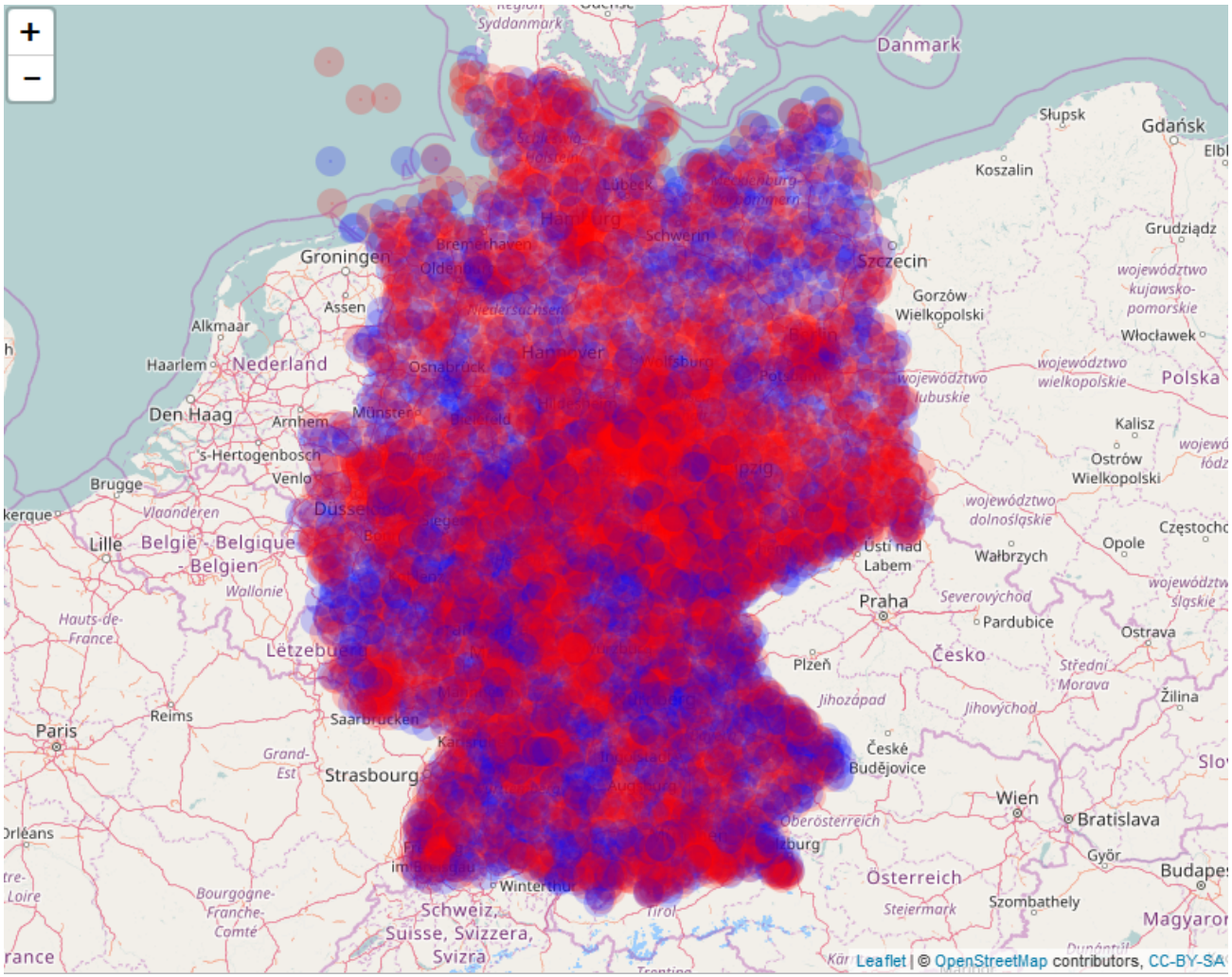
Data can be plotted with regular R code

```
plot(clim[,c(2,4)], type="l", xaxt="n", las=1)
berryFunctions::monthAxis(ym=TRUE)   ;   abline(h=0)
```

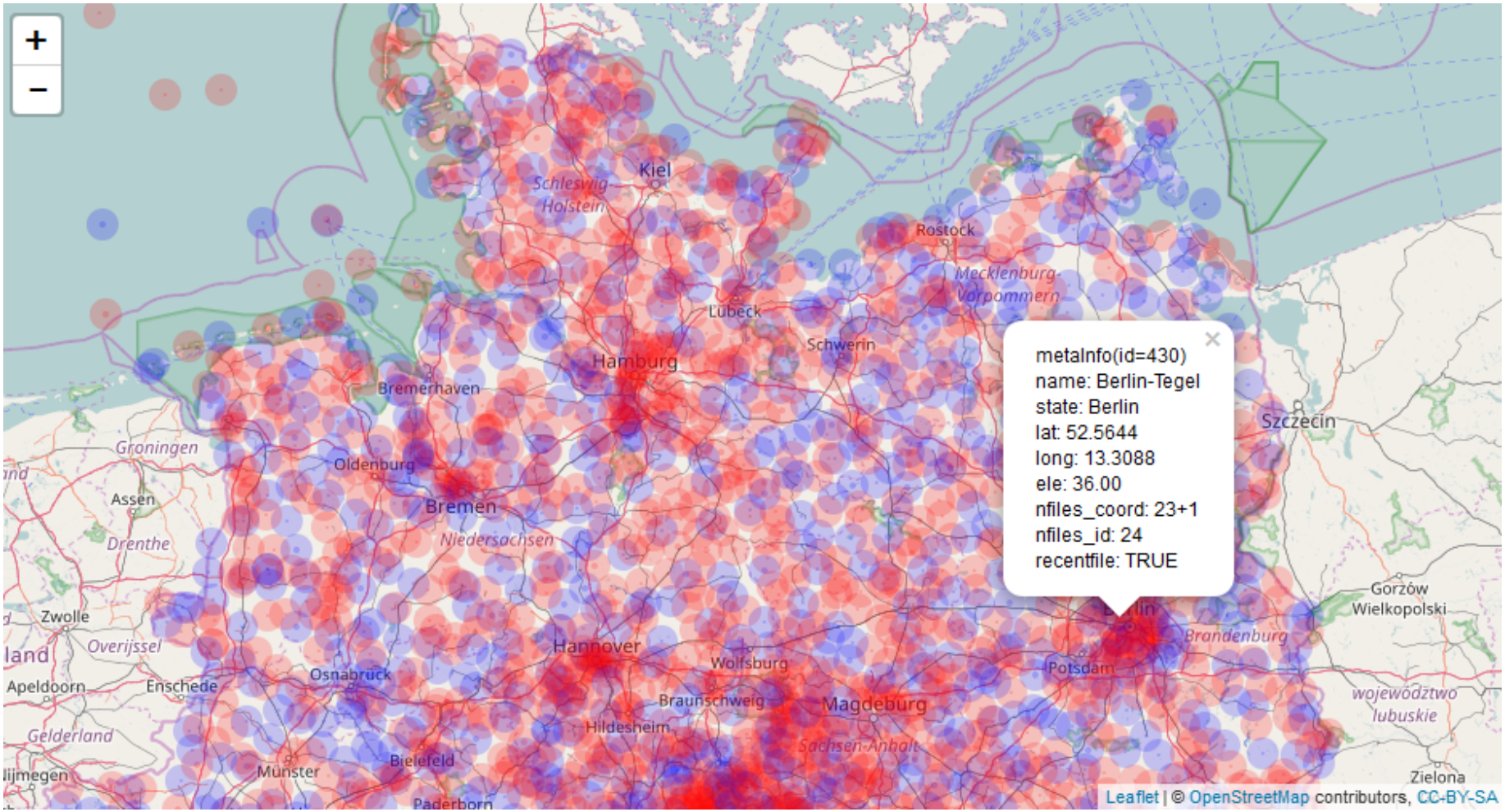


Interactive map (local)

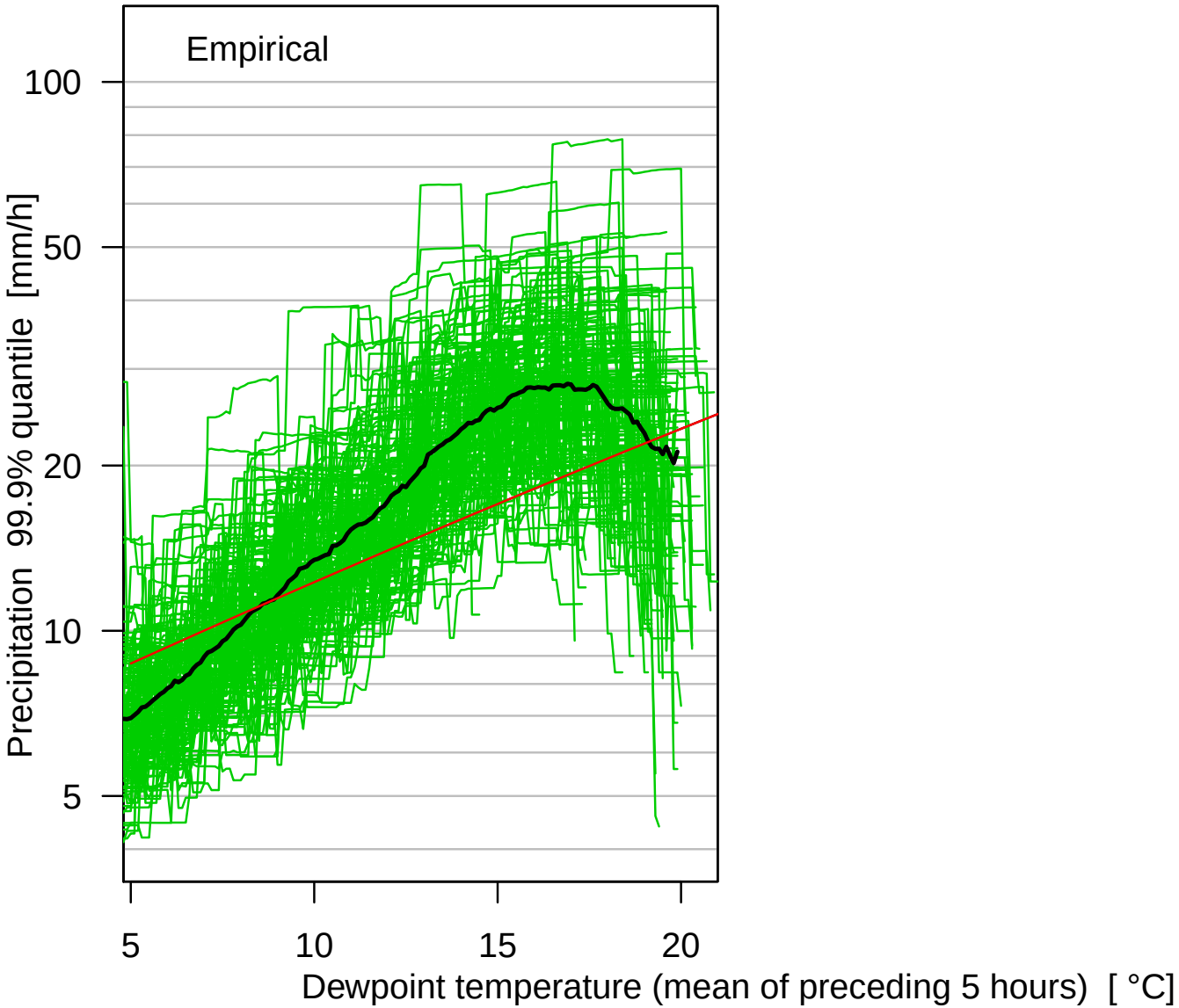
```
vignette("mapDWD", package="rdwd")
```



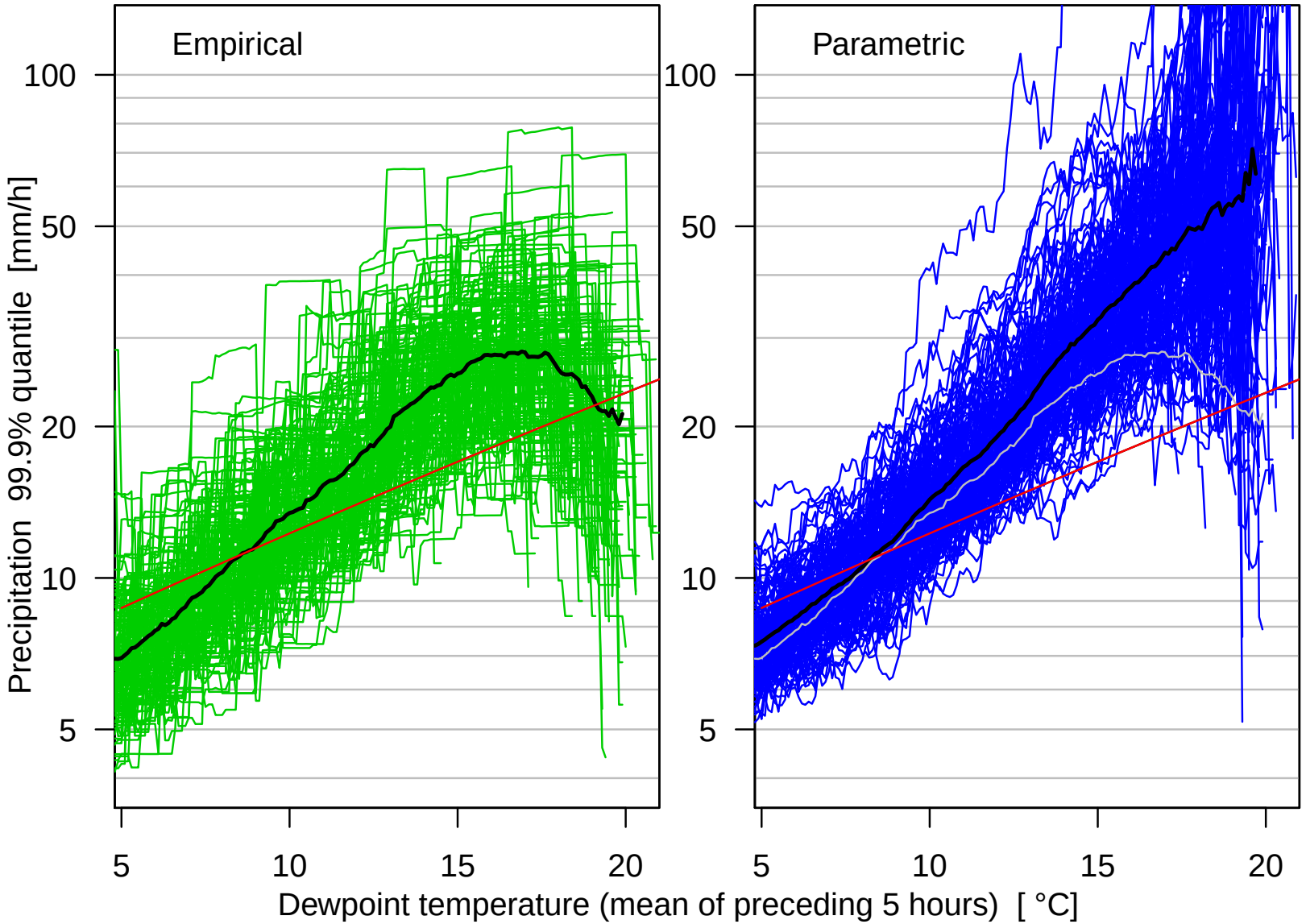
Interactive map (CRAN)



Extreme rainfall over temperature (github.com/brry/prectemp)



Extreme rainfall over temperature (github.com/brry/prectemp)



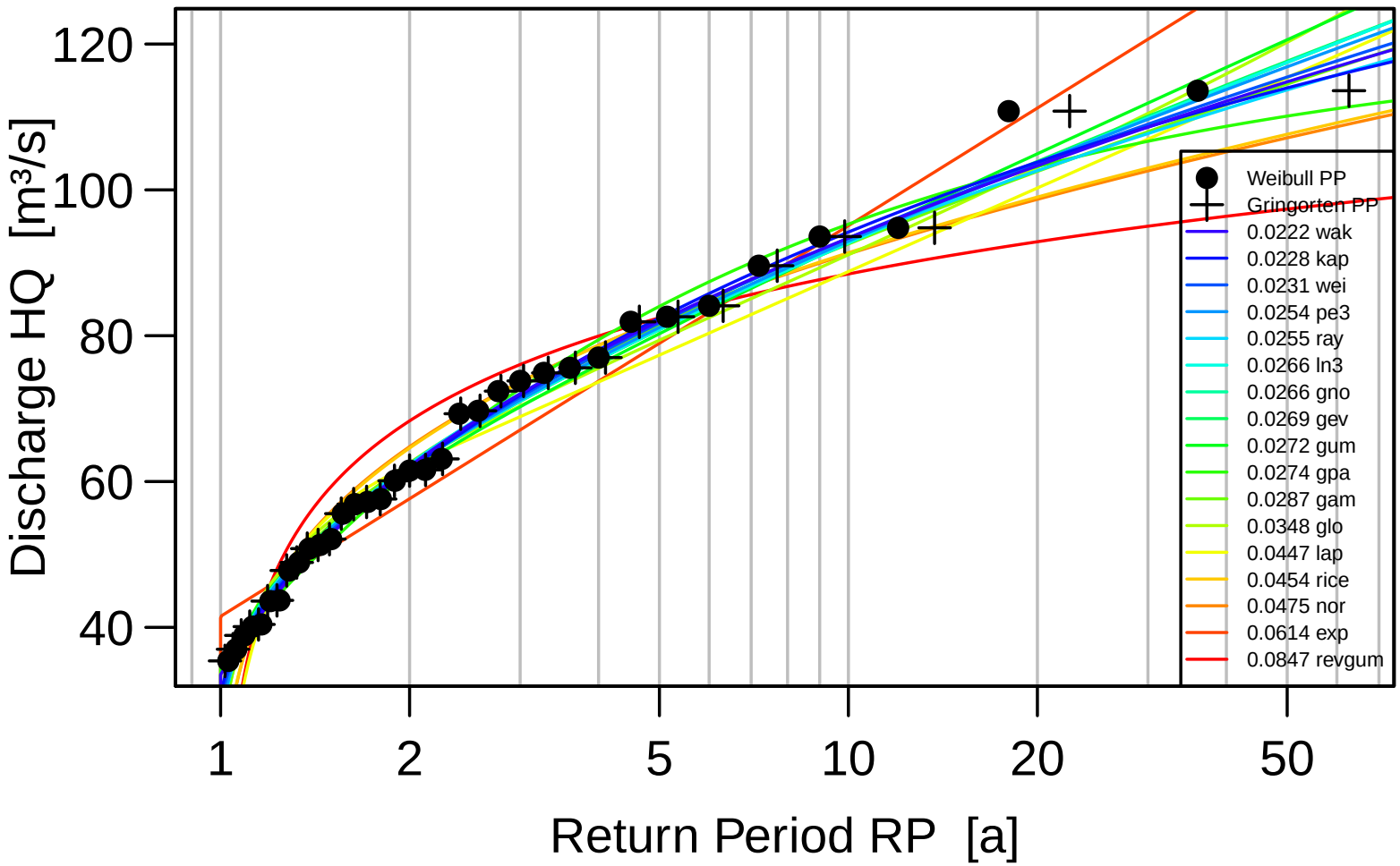
Return periods

Return periods

```
library(extremeStat) ; data(annMax)
dlf <- distLextreme(annMax, quiet=TRUE)
```

Return periods

```
library(extremeStat) ; data(annMax)
dlf <- distLextreme(annMax, quiet=TRUE)
plotLextreme(dlf, log=TRUE, nbest=17)
```

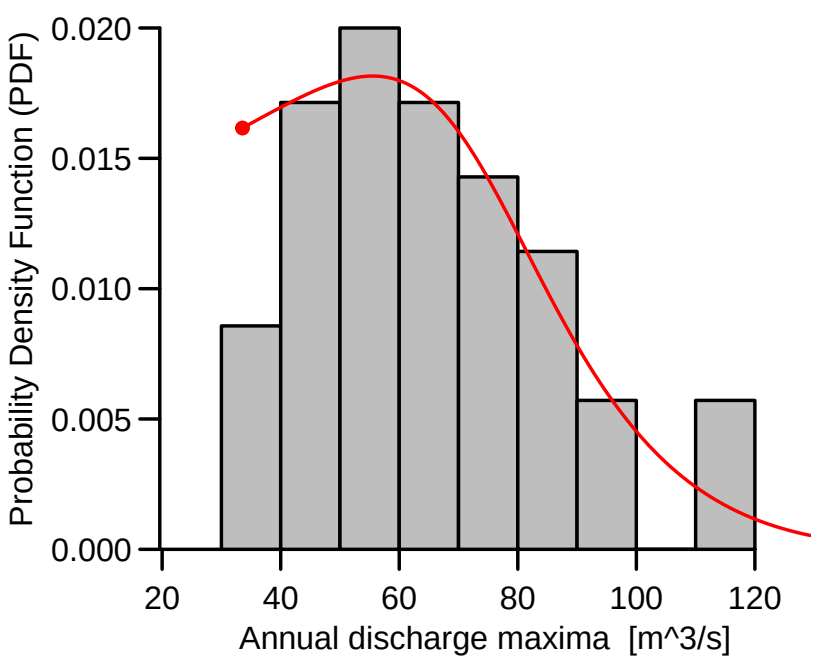


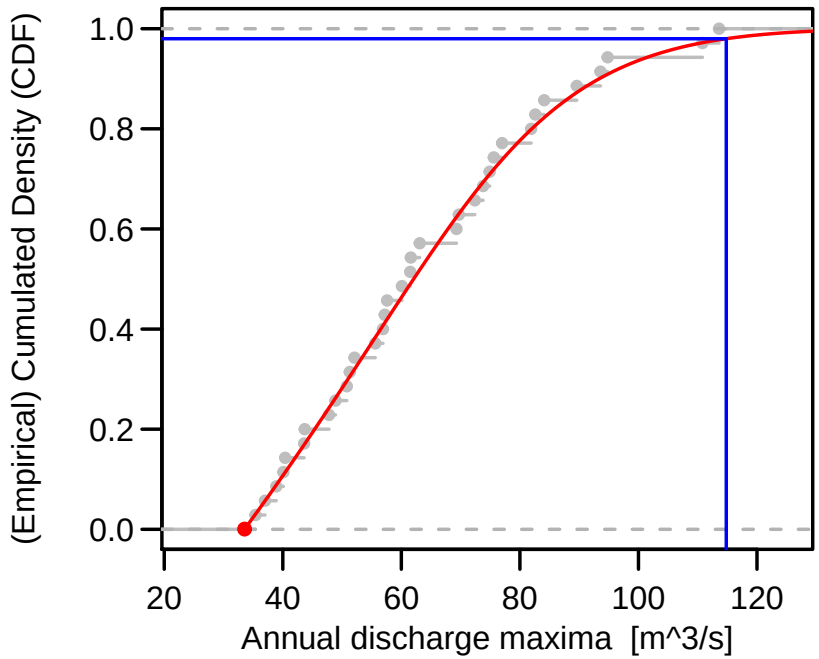
Return periods (RP) - Return Level (RL)

$$RP = \frac{1}{P_{exceedance}} = \frac{1}{1-P_{nonexceedance}} \quad e.g.\ 50 = \frac{1}{0.02} = \frac{1}{1-0.98}$$

$$P_{nonexceedance} = 1 - \frac{1}{RP} \quad e.g.\ 0.98 = 1 - \frac{1}{50}$$

$$RL = quantile\_function(probs = P_{nonexceedance},\ parameter = fitted\_par)$$



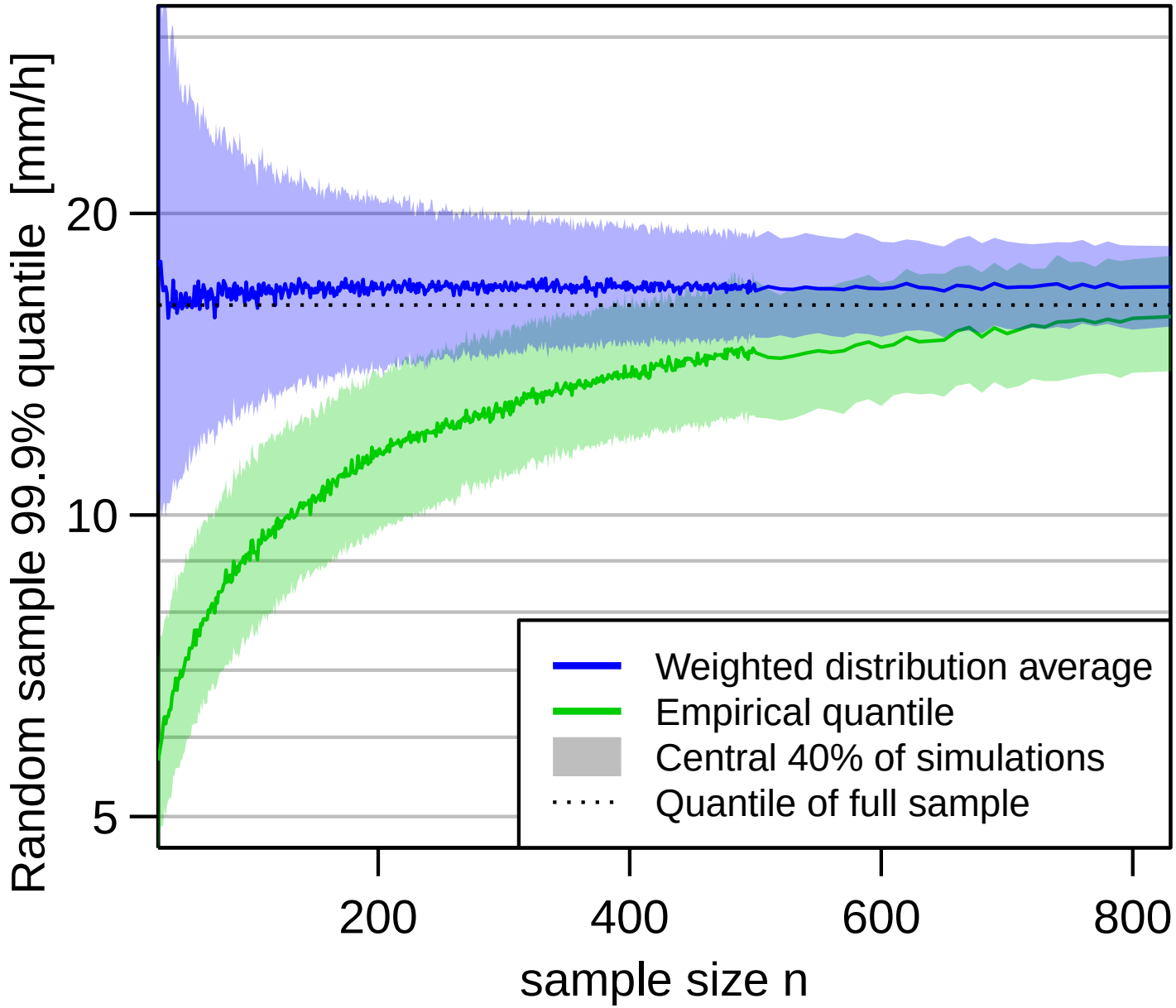



```
distLextreme(annMax, RPs=c(10,20,50,100) )$returnlev
```

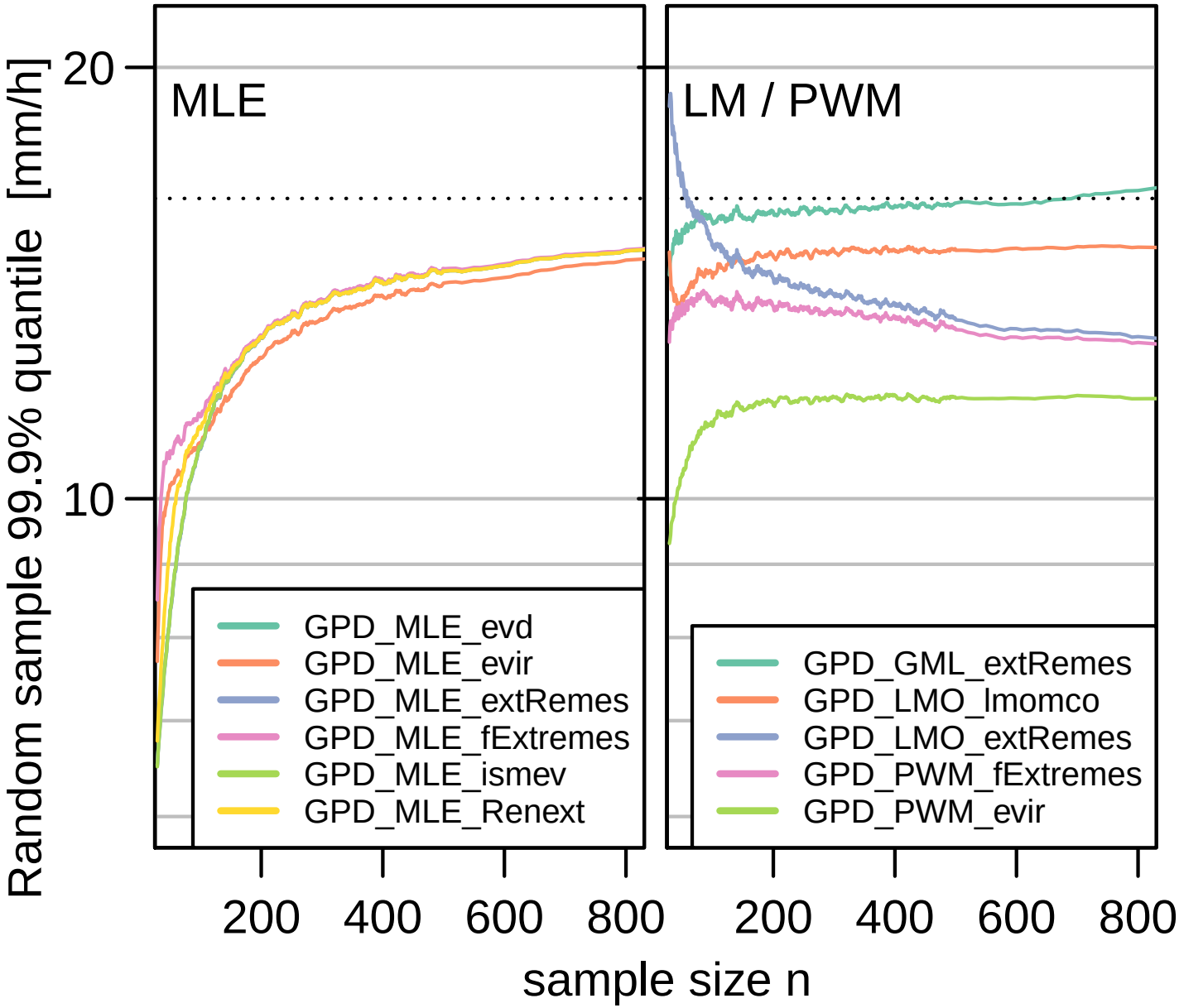
Three methods for extreme value analysis

- Empirical quantiles (order observations, take $(p \cdot n)^{th}$ value)
- BM: Block Maxima (e.g. annual maxima, fit GEV, use qGEV)
- POT: Peak Over Threshold (fit GPD to values above threshold)

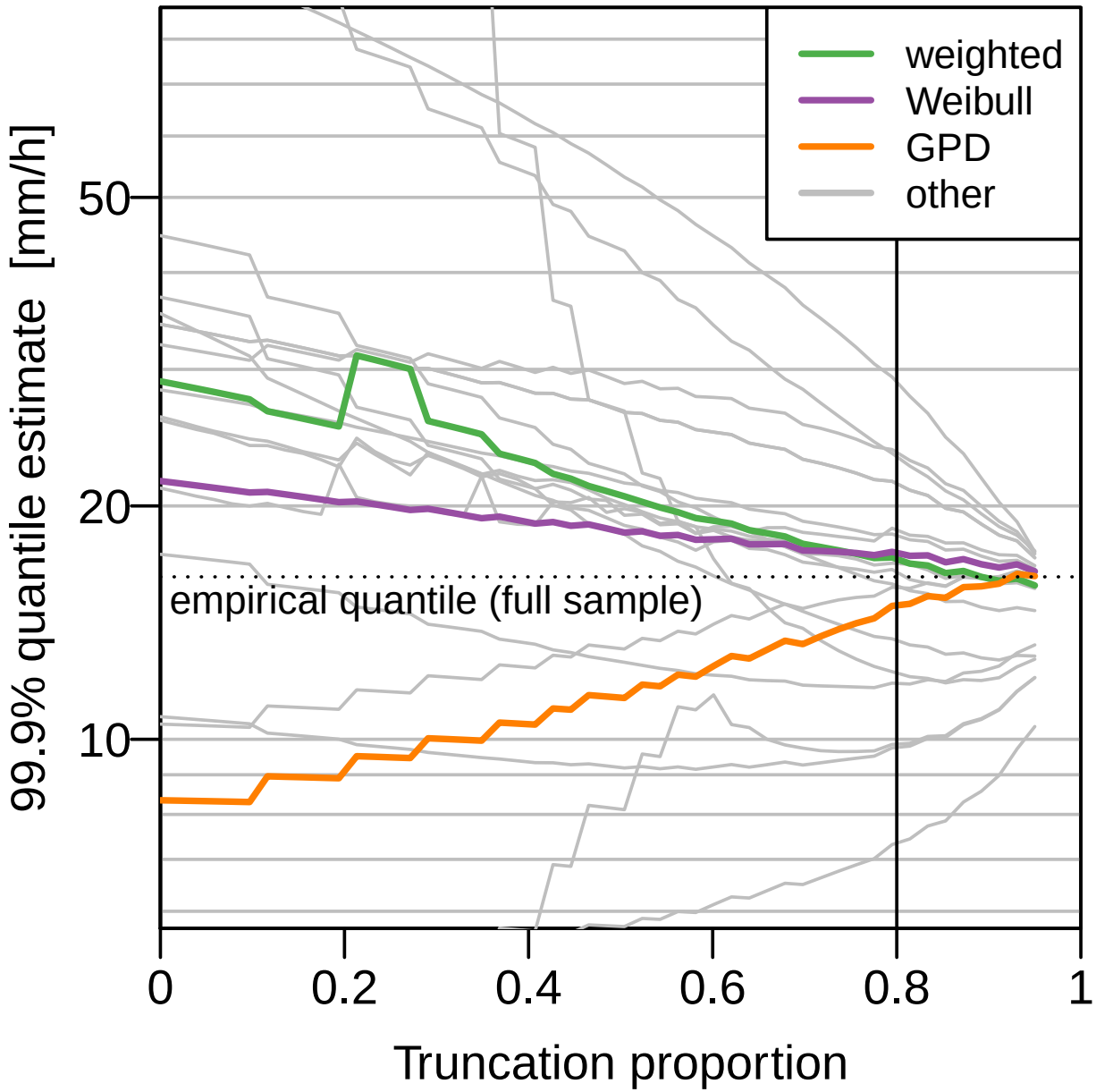
Empirical quantiles are underestimated in small samples



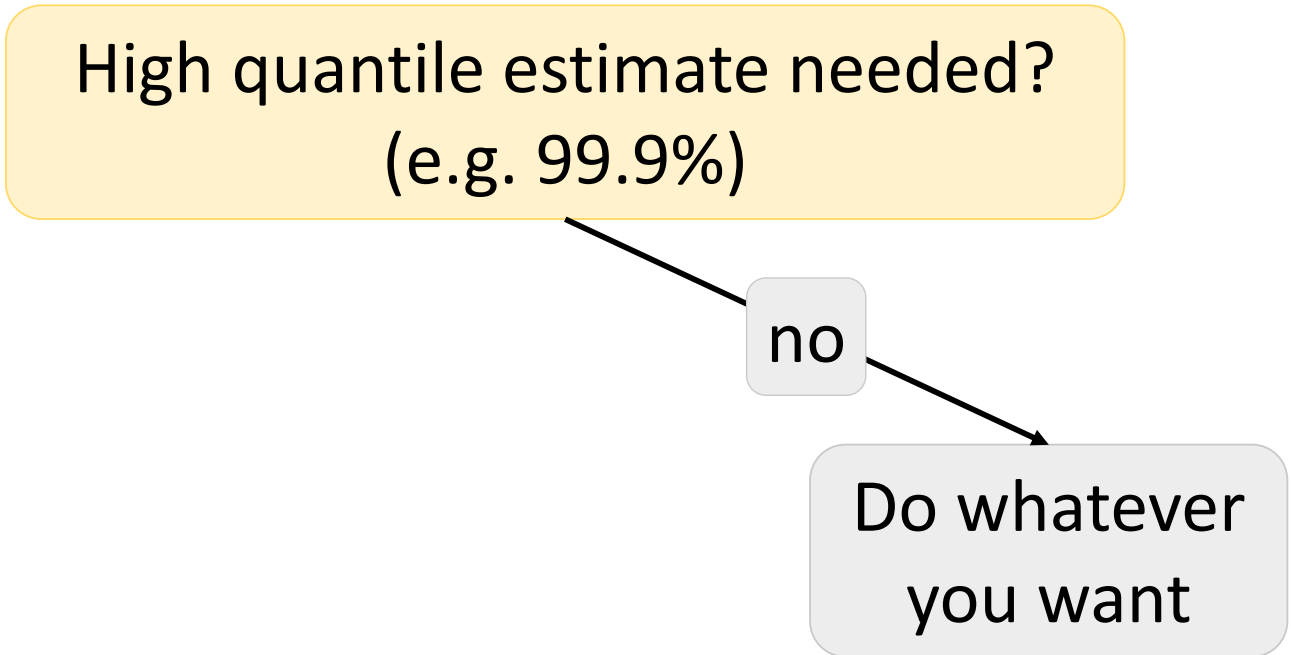
Don't use GPD with Maximum Likelihood Estimation for fitting



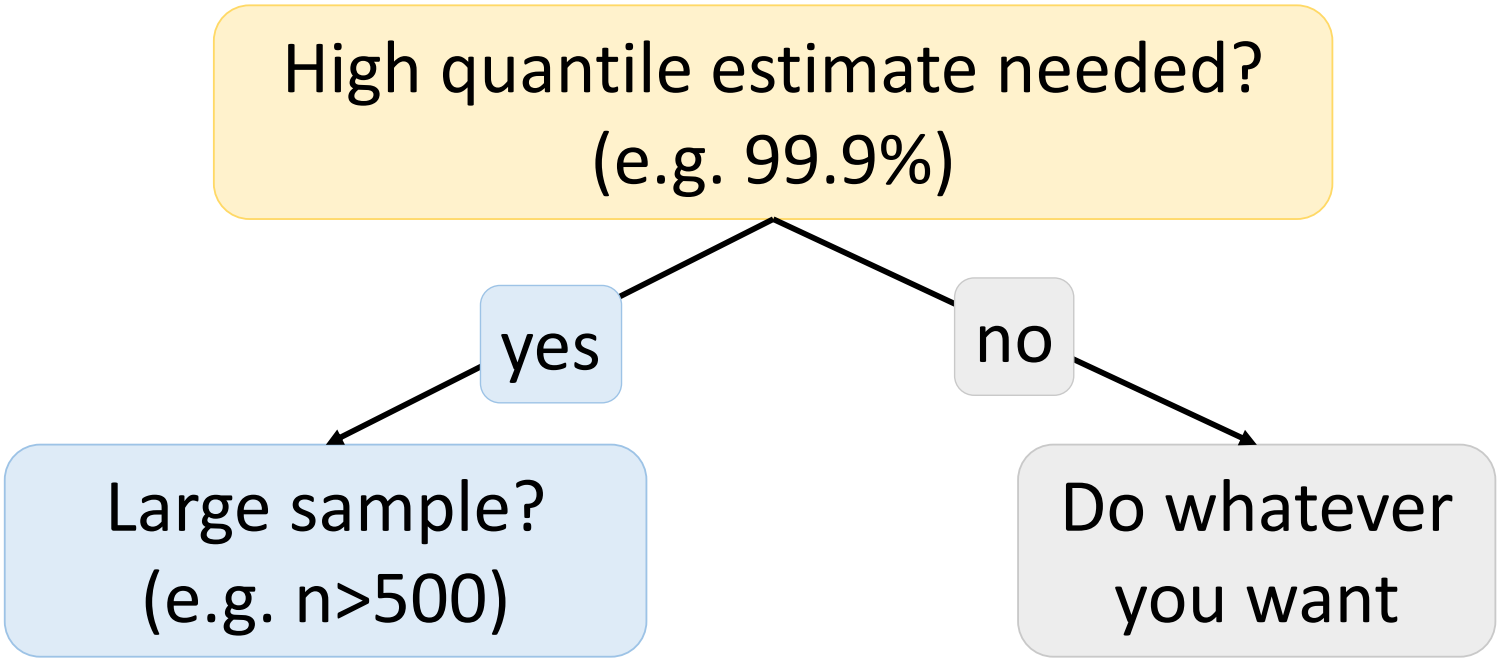
Choose threshold wisely



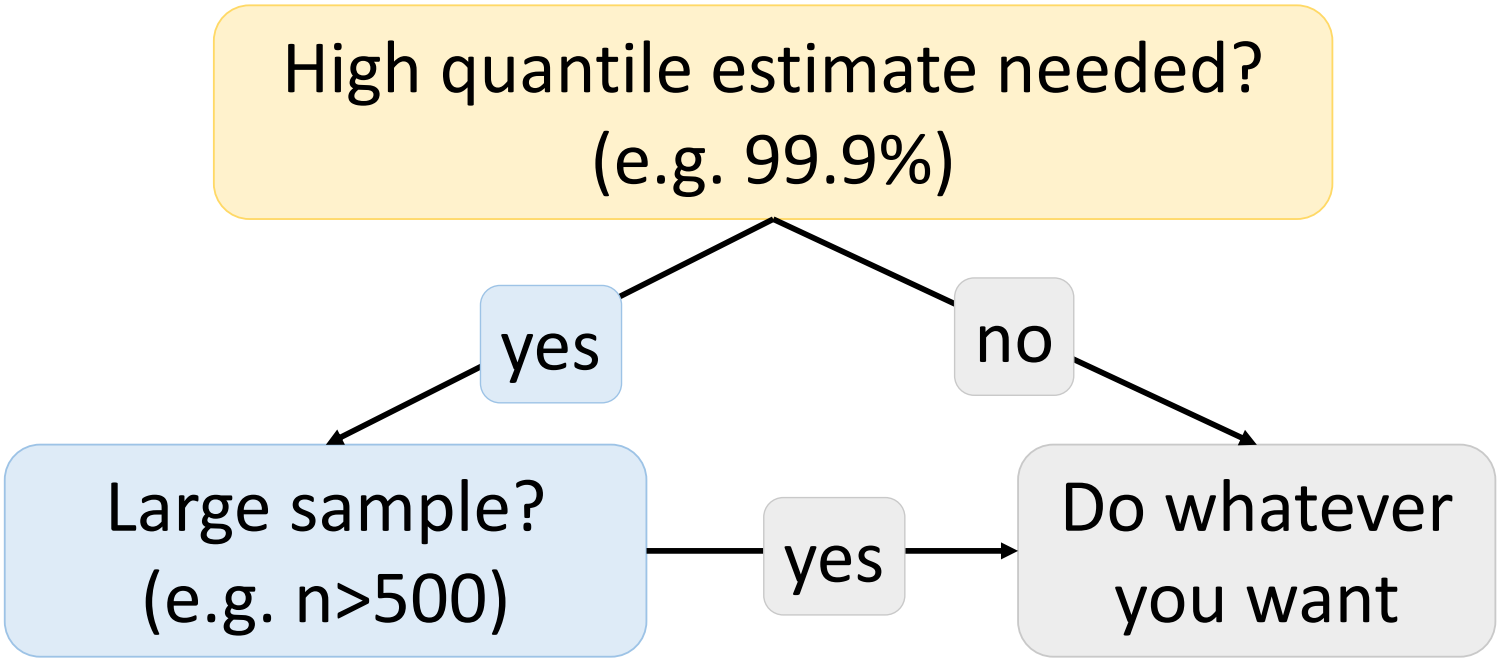
Conclusions



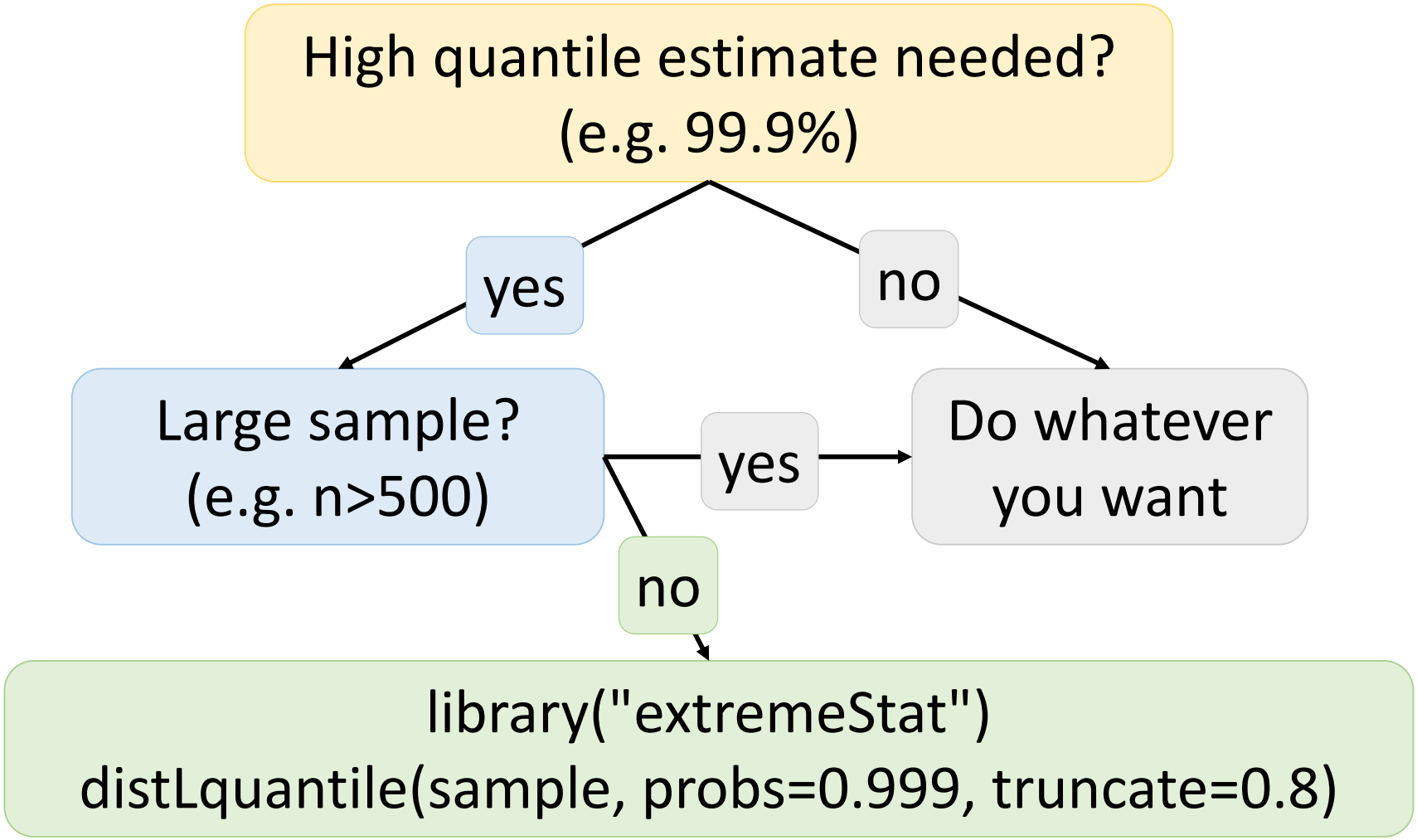
Conclusions



Conclusions



Conclusions



Conclusions

Conclusions

- DWD has tons of observational (and derived) datasets

Conclusions

- DWD has tons of observational (and derived) datasets
- Usage of the data is easy with `rdwd`

Conclusions

- DWD has tons of observational (and derived) datasets
- Usage of the data is easy with `rdwd`
- Extreme quantile estimation is hard for small samples

Conclusions

- DWD has tons of observational (and derived) datasets
- Usage of the data is easy with `rdwd`
- Extreme quantile estimation is hard for small samples
- `extremeStat` does the nasty part for you

Conclusions

- DWD has tons of observational (and derived) datasets
- Usage of the data is easy with `rdwd`
- Extreme quantile estimation is hard for small samples
- `extremeStat` does the nasty part for you

- github.com/brry/rdwd

Conclusions

- DWD has tons of observational (and derived) datasets
- Usage of the data is easy with `rdwd`
- Extreme quantile estimation is hard for small samples
- `extremeStat` does the nasty part for you

- github.com/brry/rdwd
- github.com/brry/extremeStat

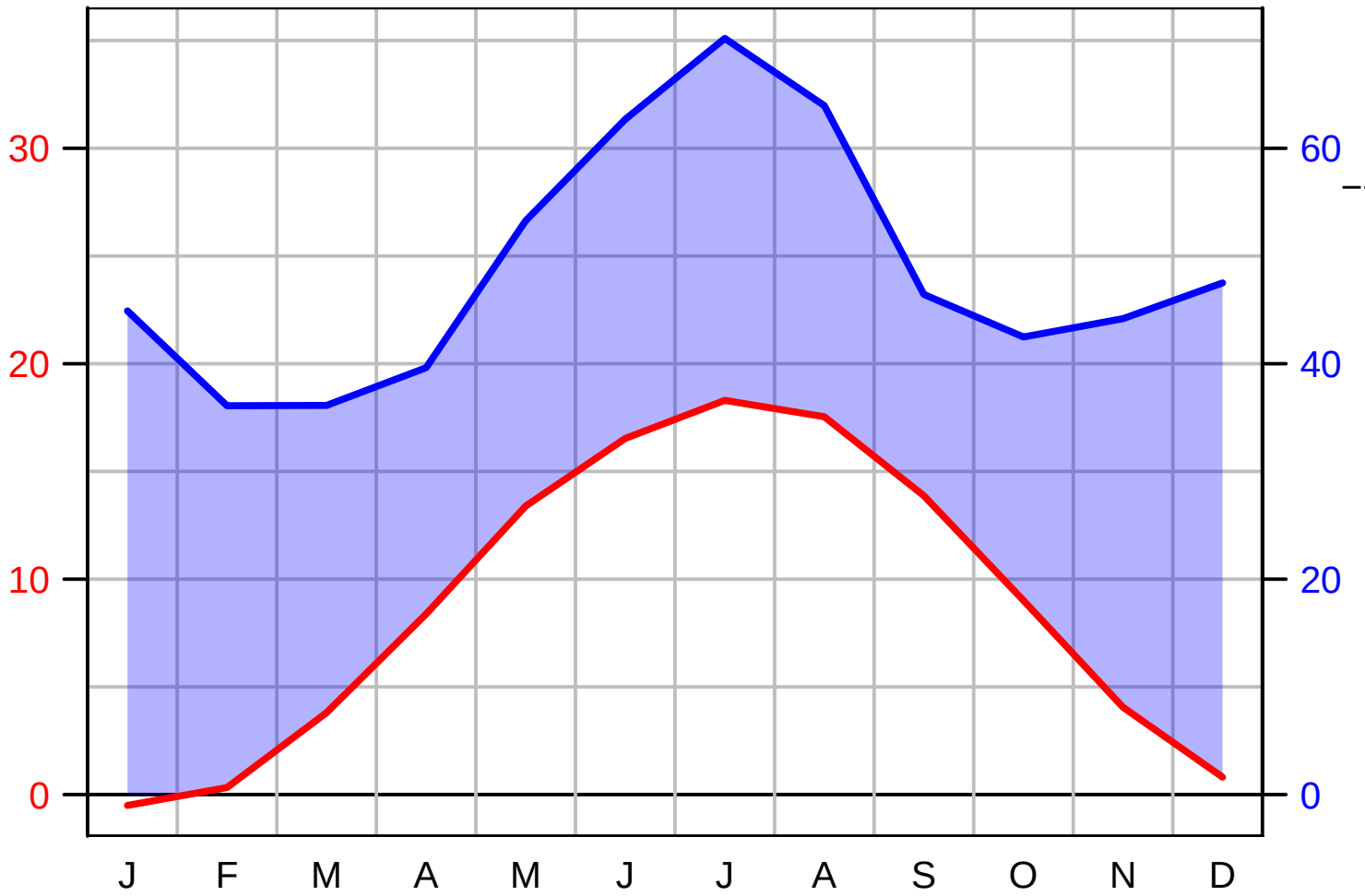
Long term climate graph (Potsdam 1893:2015)

```
clim <- dataDWD(selectDWD("Potsdam", res="monthly",
                          var="kl", per="h"))
clim$month <- substr(clim$MESS_DATUM_BEGINN,5,6)
temp <- tapply(clim$LUFTTEMPERATUR, clim$month, mean)
prec <- tapply(clim$NIEDERSCHLAGSHOEHE, clim$month, mean)
berryFunctions::climateGraph(temp, prec, main="")
```

Long term climate graph (Potsdam 1893:2015)

Ø 8.8 °C

∑ 587.5 mm

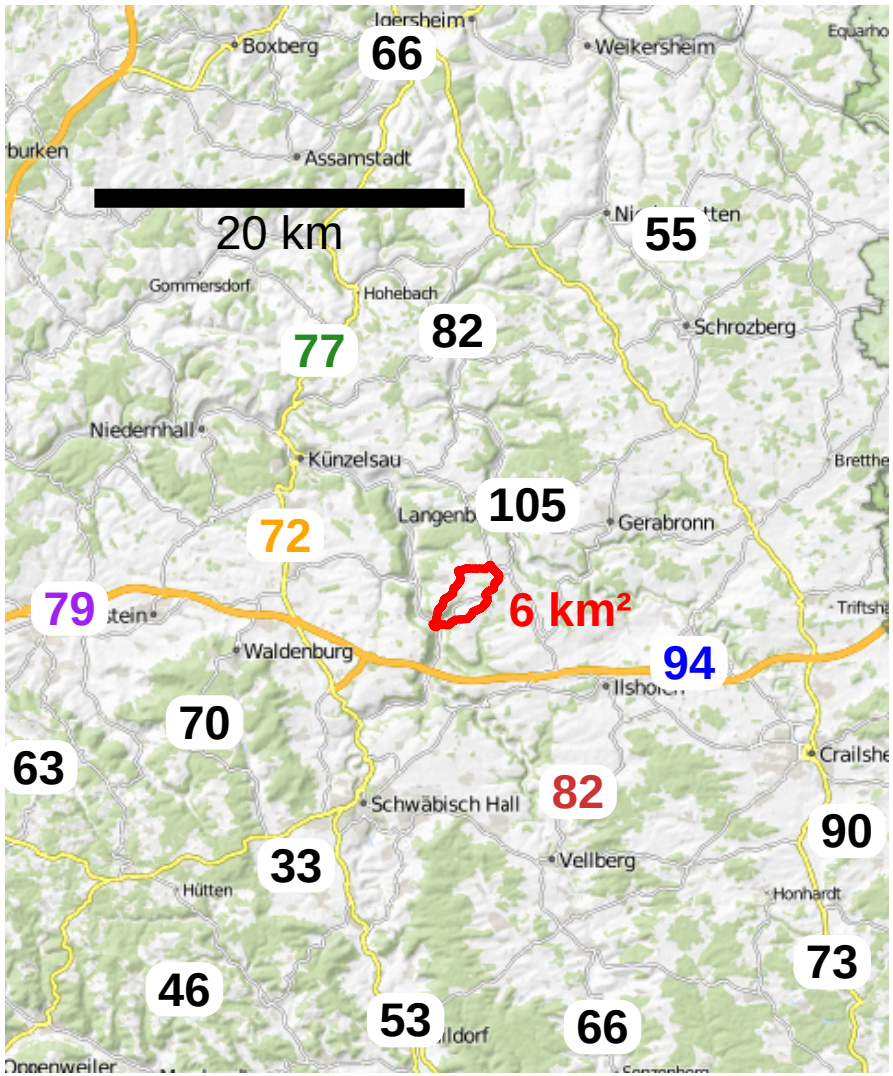


m	T °C	P mm
J	−0.5	45
F	0.3	36
M	3.8	36
A	8.4	40
M	13.4	53
J	16.5	63
J	18.3	70
A	17.5	64
S	13.9	46
O	9.0	42
N	4.1	44
D	0.8	48

Flashflood event rainfall analysis (Taskforce report)

Flashflood event rainfall analysis (Taskforce report)

Rainfall sum 2016−05−29 7AM−7AM [mm]



Temporal succession