

rdwd - an R package to select, download and read climate data from the German Weather Service

Berry Boessenkool, uni-potsdam.de, Feb 2017

`berry-b@gmx.de`

github.com/brry/rdwd

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
 taoeswerte_RR_00053_akt.zip	6 KB	27.01.2017	06:33:00

R saves the day

R package `rdwd` —> easy usage of the datasets

Overview

- ▶ Motivation
- ▶ Usage
- ▶ Applications
- ▶ Community

Usage

- get URL
- download
 - read
 - plot
 - map

U1/5: Get dataset URL with `selectDWD`

U1/5: Get dataset URL with selectDWD

```
library("rdwd")
```

U1/5: Get dataset URL with selectDWD

```
library("rdwd")
```

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

U1/5: Get dataset URL with selectDWD

```
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
```

U2/5: Download dataset with dataDWD

U2/5: Download dataset with dataDWD

```
file <- dataDWD(link, read=FALSE)
```

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

U2/5: Download dataset with dataDWD

```
file <- dataDWD(link, read=FALSE)
```

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

```
file
```

```
## [1] "daily_kl_recent_tageswerte_KL_03987_akt.zip"
```

U3/5: Unzip file and read + convert data with readDWD

U3/5: Unzip file and read + convert data with readDWD

```
clim <- readDWD(file)
```


U3/5: Unzip file and read + convert data with readDWD

```
clim <- readDWD(file)
```

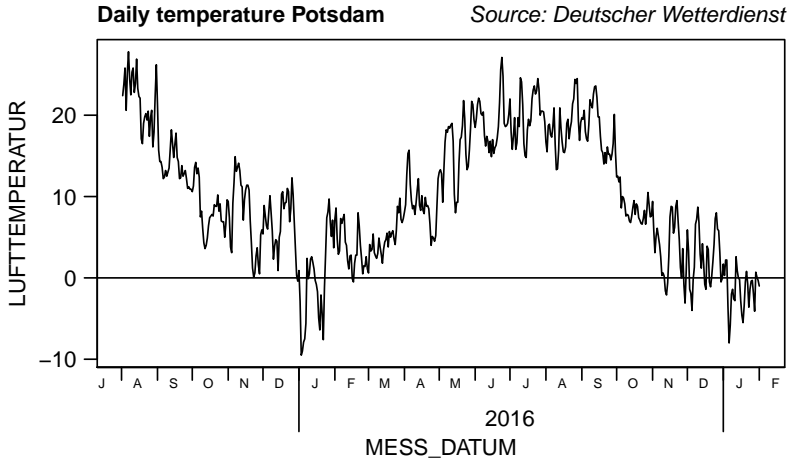
```
str(clim)
```

```
## 'data.frame': 550 obs. of 18 variables:
## $ STATIONS_ID : int 3987 3987 3987 3987 3987 3987 3987 3987 3987 3987 ...
## $ MESS_DATUM : POSIXct, format: "2015-08-02" "2015-08-03" ...
## $ QUALITAETS_NIVEAU : int 3 3 3 3 3 3 3 3 3 3 ...
## $ LUFTTEMPERATUR : num 22.4 23.8 25.8 20.6 25.2 27.8 24.5 22.5 25.3 25.8 ...
## $ DAMPDRUCK : num 11.7 13.3 15.7 15.4 15.8 17.4 18.6 15.3 17.7 19.1 ...
## $ BEDECKUNGSGRAD : num 4.5 2 2.9 4.9 3.6 3.4 4.4 2.3 3.3 4 ...
## $ LUFTDRUCK_STATIONSHOEHE : num 1007 1006 1002 1007 1004 ...
## $ REL_FEUCHTE : num 46.8 49.4 52.4 66.2 54.1 ...
## $ WINDGESCHWINDIGKEIT : num 3 3 5 3.4 3.4 4 4.3 3.5 3.8 3.8 ...
## $ LUFTTEMPERATUR_MAXIMUM : num 30 32.3 35.3 26.3 34.6 37.6 33.3 29.5 33.5 33.1 ...
## $ LUFTTEMPERATUR_MINIMUM : num 15.1 14 18.4 16.4 16.1 21.2 19.2 16.6 17.1 19.4 ...
## $ LUFTTEMP_AM_ERDB_MINIMUM : num 11.6 11.7 16.1 14.9 13.5 18.1 17.8 15.5 16.1 18.7 ...
## $ WINDSPITZE_MAXIMUM : num 8.1 9.2 17.3 9.1 9.6 9.1 12.5 8.2 8.4 11.7 ...
## $ NIEDERSCHLAGSHOEHE : num 0 0 4.1 0 0 0 0.1 0 0 0 ...
## $ NIEDERSCHLAGSHOEHE_IND : int 0 0 6 0 0 0 6 0 0 0 ...
## $ SONNENSCHEINDAUER : num 13.4 14.4 11.6 10.7 13.3 ...
## $ 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 ...
```

U4/5: Data can be plotted with regular R code

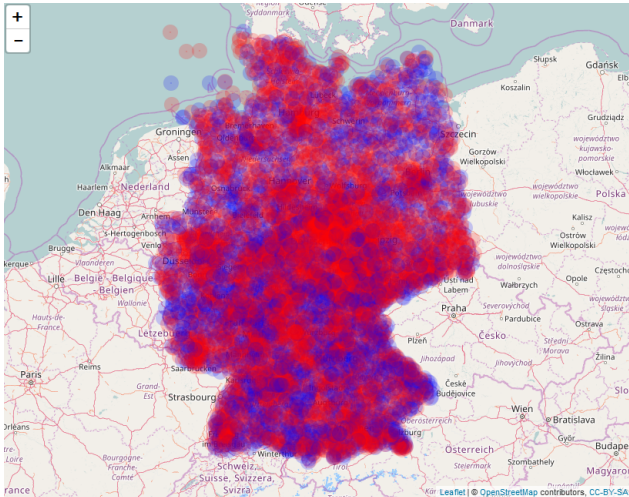
U4/5: Data can be plotted with regular R code

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

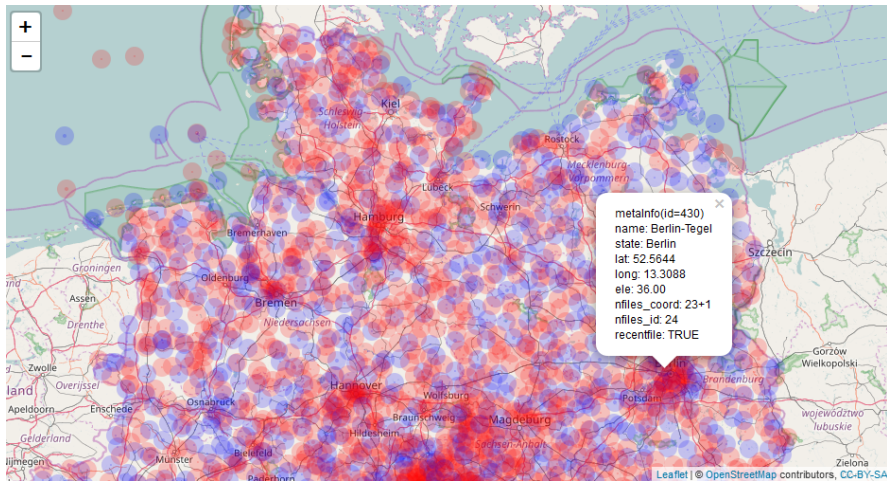


U5/5: Interactive map (mapDWD.html)

```
data(mapDWD) ; library(leaflet) ; mapDWD
```



U5/5: Interactive map (mapDWD.html)



Applications

- climate graph
- event analysis
- rainfall extremes

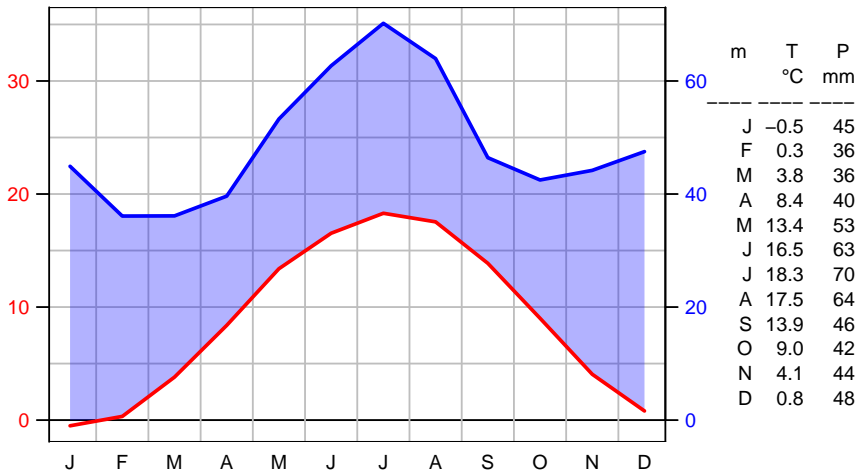
A1/3: Long term climate graph (Potsdam 1893:2015)

```
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="")
```

A1/3: Long term climate graph (Potsdam 1893:2015)

Ø 8.8 °C

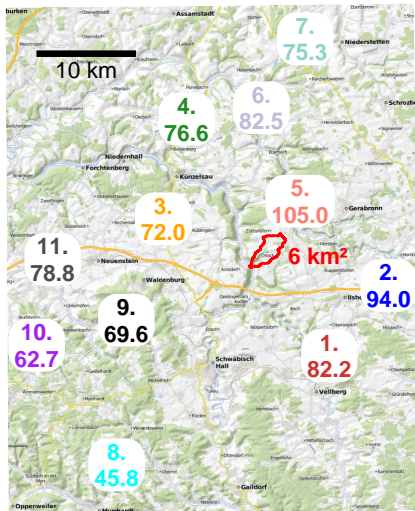
Σ 587.5 mm



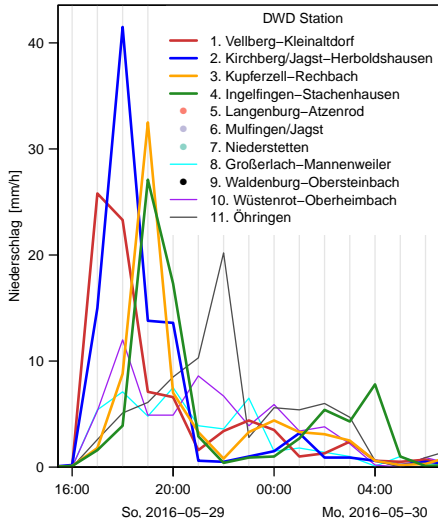
A2/3: Flashflood event rainfall analysis (Taskforce report)

A2/3: Flashflood event rainfall analysis (Taskforce report)

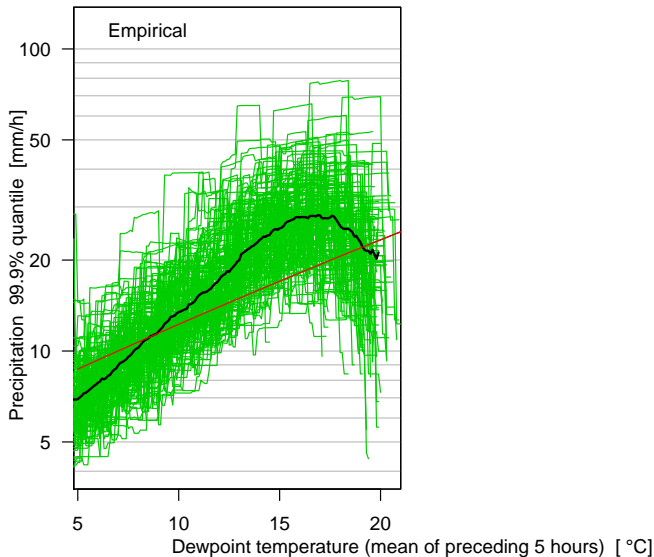
DWD–Stationen bei Braunsbach



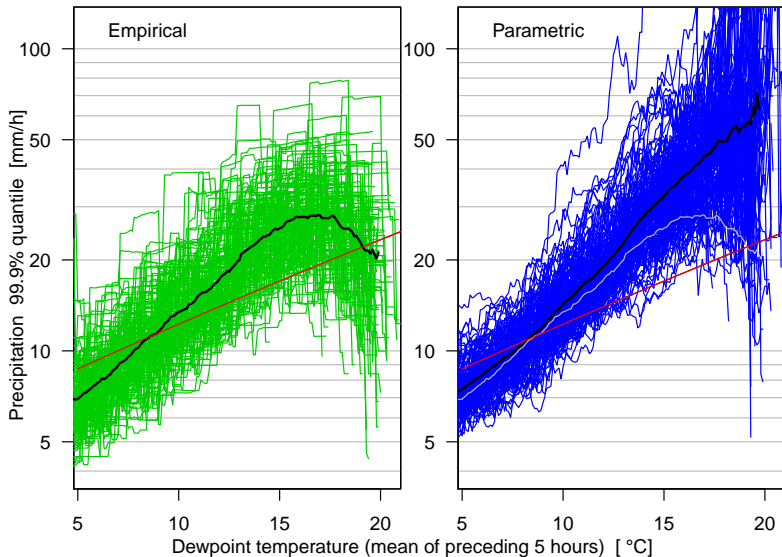
Niederschlagsverlauf



A3/3: Extreme rainfall over temperature (github.com/brry/prectemp)



A3/3: Extreme rainfall over temperature (github.com/brry/prectemp)



The FOSS community role

The FOSS community role

- ▶ Stackoverflow for programming help

The FOSS community role

- ▶ Stackoverflow for programming help
- ▶ Lobbying DWD into publishing tax-paid data

The FOSS community role

- ▶ Stackoverflow for programming help
- ▶ Lobbying DWD into publishing tax-paid data
- ▶ Package distribution infrastructure (CRAN)

The FOSS community role

- ▶ Stackoverflow for programming help
- ▶ Lobbying DWD into publishing tax-paid data
- ▶ Package distribution infrastructure (CRAN)
- ▶ `leaflet` interactive map really easy to create

Conclusion

Conclusion

- ▶ FOSS is awesome

Conclusion

- ▶ FOSS is awesome
- ▶ DWD is awesome

Conclusion

- ▶ FOSS is awesome
- ▶ DWD is awesome
- ▶ Usage of the data is easy with `rdwd`