Weather and Soil Data with R: Retrieval and Visualization

Fernando Miguez

11/01/2022

Abstract

This short workshop will be an introduction to using R for retrieval and visualization of weather and soil data. We will use the *apsimx* R package and perform simple analyses and interpretation through the use of tables and figures. It is recommended that you have R installed (version 4.0.0 or newer) and several other R packages. Previous familiarity with R is desirable.

Preliminaries

Install the followig packages ahead of time using the following command.

```
install.packages(c("apsimx", "nasapower", "daymetr", "GSODR", "soilDB", "ggplot2", "maps", "sp", "sf",
```

```
## Load libraries
library(apsimx)
library(ggplot2)
```

Weather data

There are many different types of weather data. For use in agricultural applications, two relevant types are station data and gridded data. Station data are derived from weather stations and might have different levels of quality control. Gridded data is normally derived from a climate/weather model and it is normally a combination of observations and models

Station data

We will first look at station data. The original source of the data is the Iowa Environmental Mesonet (IEM: https://mesonet.agron.iastate.edu/). The IEM provides station-level data, but the solar radiation is derived from NOAA (https://rapidrefresh.noaa.gov/hrrr/), so it is not observed.

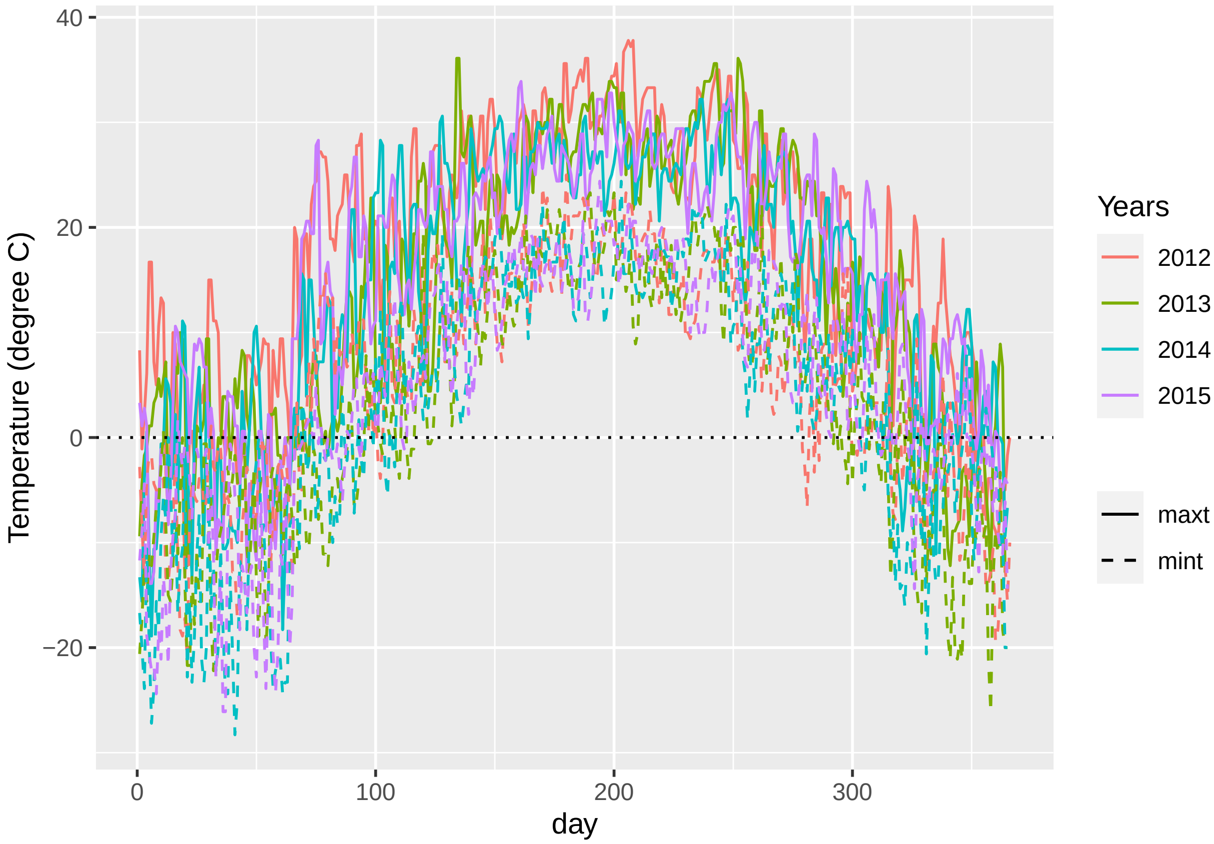
```
## This code 'gets' the data through a call to IEM
ames.iem <- get_iem_apsim_met(lonlat = c(-93.77, 42.02),
                              dates = c("1990-01-01", "2021-12-31"))
## There is a summary function which provides some summaries per year
summary(ames.iem)
```

```
##    year months days high_maxt high_mint avg_maxt avg_mint low_maxt low_mint
## 1  1990   1:12 1:31      35.6      24.4    16.26     4.47    -21.7    -26.1
## 2  1991   1:12 1:31      35.6      23.9    15.19     4.10    -13.9    -25.6
## 3  1992   1:12 1:31      31.7      22.2    14.75     3.95    -12.2    -23.9
```

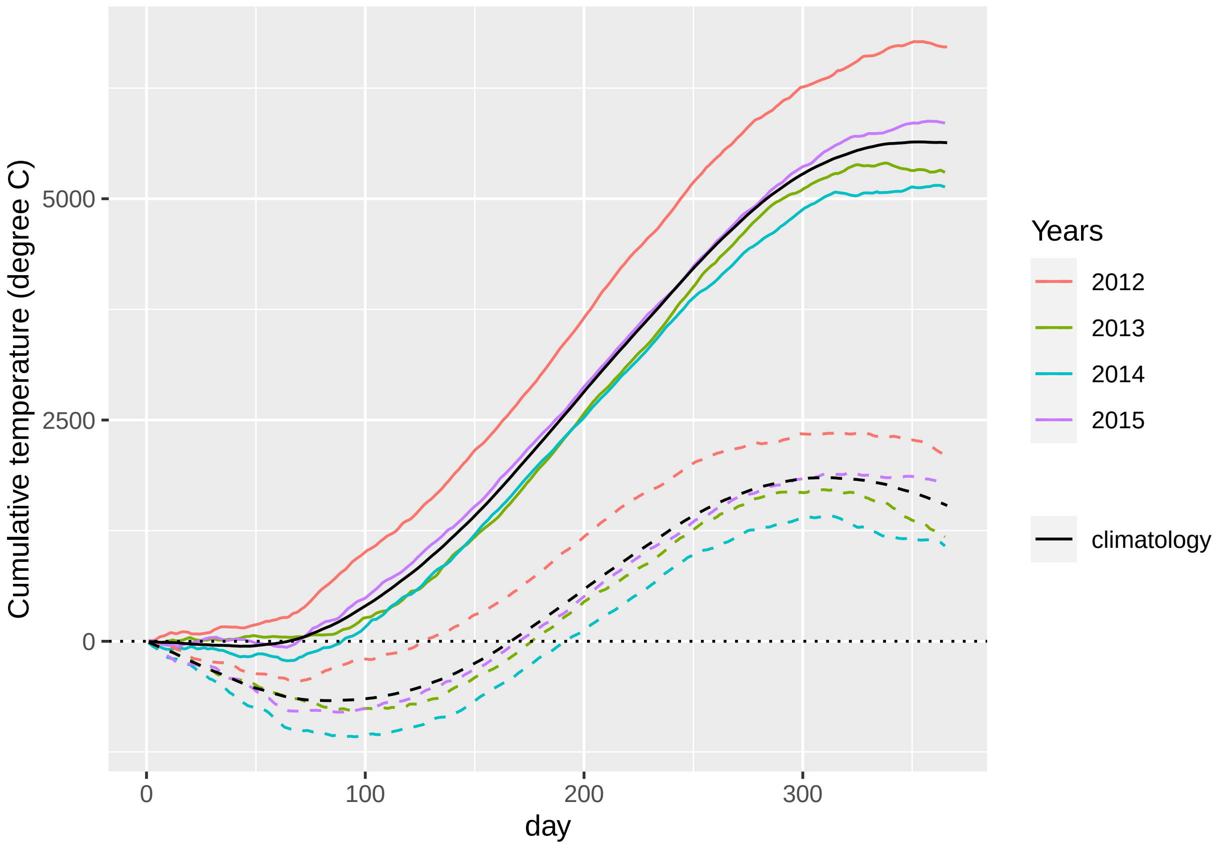
```
## 4  1993   1:12 1:31      31.7      21.7    12.84     3.31    -14.4    -24.4
## 5  1994   1:12 1:31      33.3      23.3    14.98     3.34    -22.8    -31.1
## 6  1995   1:12 1:31      36.1      23.9    14.19     3.57    -16.1    -23.3
## 7  1996   1:12 1:31      33.9      24.4    13.35     2.39    -23.3    -33.3
## 8  1997   1:12 1:31      35.0      23.3    14.31     3.59    -17.8    -25.6
## 9  1998   1:12 1:31      33.9      22.8    15.97     5.79    -13.3    -25.0
## 10 1999   1:12 1:31      35.0      23.9    16.17     4.75    -18.9    -30.0
## 11 2000   1:12 1:31      36.7      23.3    15.82     3.97    -18.3    -26.1
## 12 2001   1:12 1:31      36.1      25.0    15.70     4.74    -14.4    -23.9
## 13 2002   1:12 1:31      35.0      23.3    16.35     4.23    -10.0    -22.8
## 14 2003   1:12 1:31      36.1      23.3    15.61     3.56    -13.3    -25.0
## 15 2004   1:12 1:31      31.7      22.8    15.60     4.29    -20.0    -25.0
## 16 2005   1:12 1:31      34.4      24.4    16.01     4.59    -16.7    -25.6
## 17 2006   1:12 1:31      33.9      23.9    17.01     5.18     -8.3    -25.0
## 18 2007   1:12 1:31      34.4      23.9    15.67     4.52    -18.3    -26.1
## 19 2008   1:12 1:31      34.4      23.9    13.72     2.93    -16.1    -27.2
## 20 2009   1:12 1:31      34.4      21.7    14.33     3.33    -20.6    -31.7
## 21 2010   1:12 1:31      33.3      23.3    15.12     4.13    -17.2    -29.4
## 22 2011   1:12 1:31      35.6      24.4    15.55     4.45    -13.9    -24.4
## 23 2012   1:12 1:31      37.8      25.0    18.35     5.72    -13.3    -21.1
## 24 2013   1:12 1:31      36.1      23.3    14.52     3.23    -13.3    -26.1
## 25 2014   1:12 1:31      32.2      24.4    14.07     2.95    -18.9    -28.3
## 26 2015   1:12 1:31      33.9      24.4    16.04     4.84    -13.9    -26.1
## 27 2016   1:12 1:31      34.4      24.4    16.76     5.84    -16.1    -27.2
## 28 2017   1:12 1:31      34.4      25.0    16.55     5.35    -20.6    -26.1
## 29 2018   1:12 1:31      36.1      23.3    14.90     4.29    -21.7    -30.0
## 30 2019   1:12 1:31      34.4      23.9    14.63     3.60    -20.6    -30.6
## 31 2020   1:12 1:31      34.4      24.4    16.27     4.51    -16.1    -24.4
## 32 2021   1:12 1:31      38.3      25.0    17.03     5.54    -21.7    -32.2
##    rain_sum radn_sum radn_avg
## 1   1125.28  6633.60    18.17
## 2   1015.75  6535.95    17.91
## 3    868.40  6502.96    17.77
## 4   1433.34  6223.47    17.05
## 5    719.34  6790.23    18.60
## 6    730.01  6391.54    17.51
## 7    935.25  6560.80    17.93
## 8    667.72  6566.43    17.99
## 9    889.72  6478.26    17.75
## 10  1017.48  6867.75    18.82
## 11   586.94  6733.54    18.40
## 12   806.94  6575.02    18.01
## 13   826.50  6637.42    18.18
## 14   883.69  6775.62    18.56
## 15   852.45  6743.66    18.43
## 16   880.12  6753.34    18.50
## 17   937.47  6694.95    18.34
## 18  1023.62  6945.62    19.03
## 19  1268.44  6745.29    18.43
## 20   938.27  6557.84    17.97
## 21  1279.89  6853.63    18.78
## 22   811.26  6763.79    18.53
## 23   641.09  7041.42    19.24
## 24   689.07  6612.84    18.12
```

```
## 25  1012.43  6653.79    18.23
## 26  1132.81  6886.00    18.87
## 27   954.53  6914.78    18.89
## 28   755.41  6794.75    18.62
## 29  1264.17  6727.35    18.43
## 30   916.69  6365.31    17.44
## 31   585.75  6880.47    18.80
## 32   627.85  6964.74    19.08
```

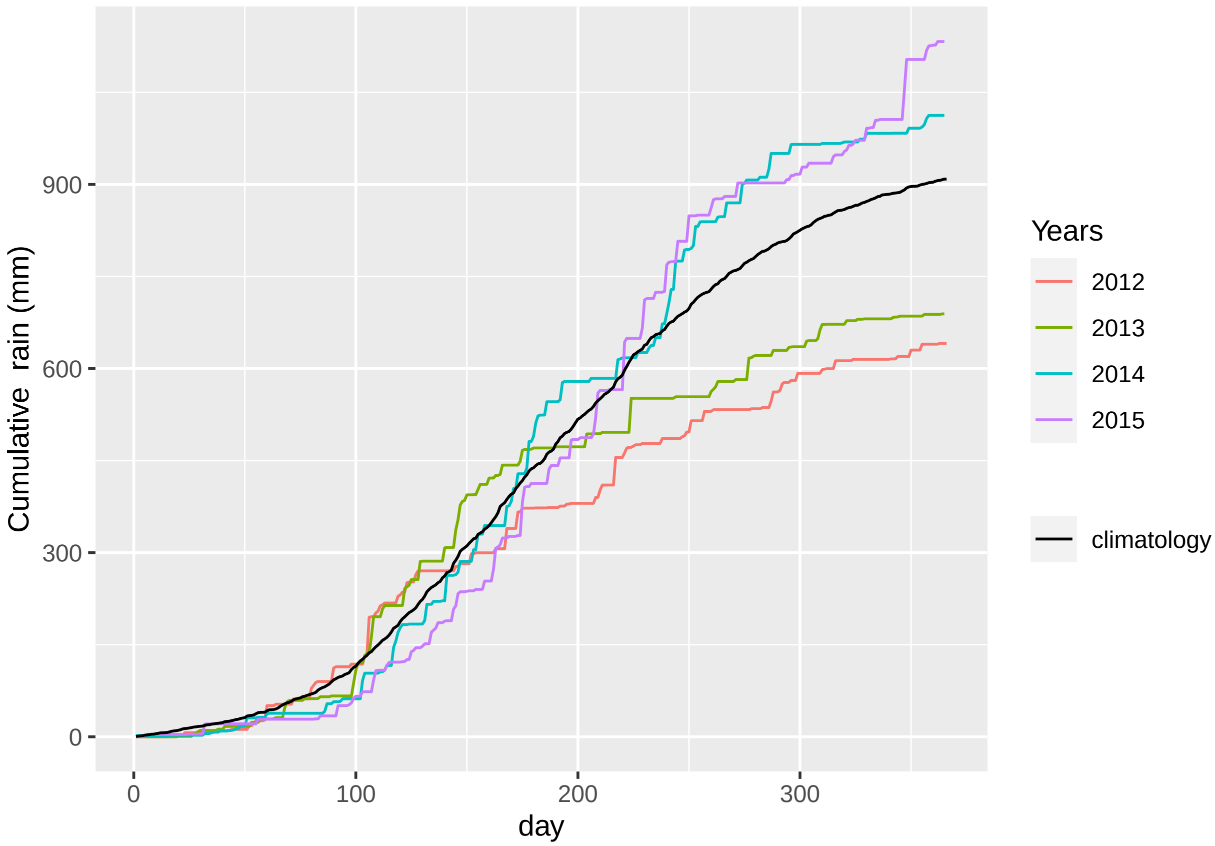
```r
## Quick visualization for just a few years, but it is still hard to see
plot(ames.iem, years = 2012:2015)
```



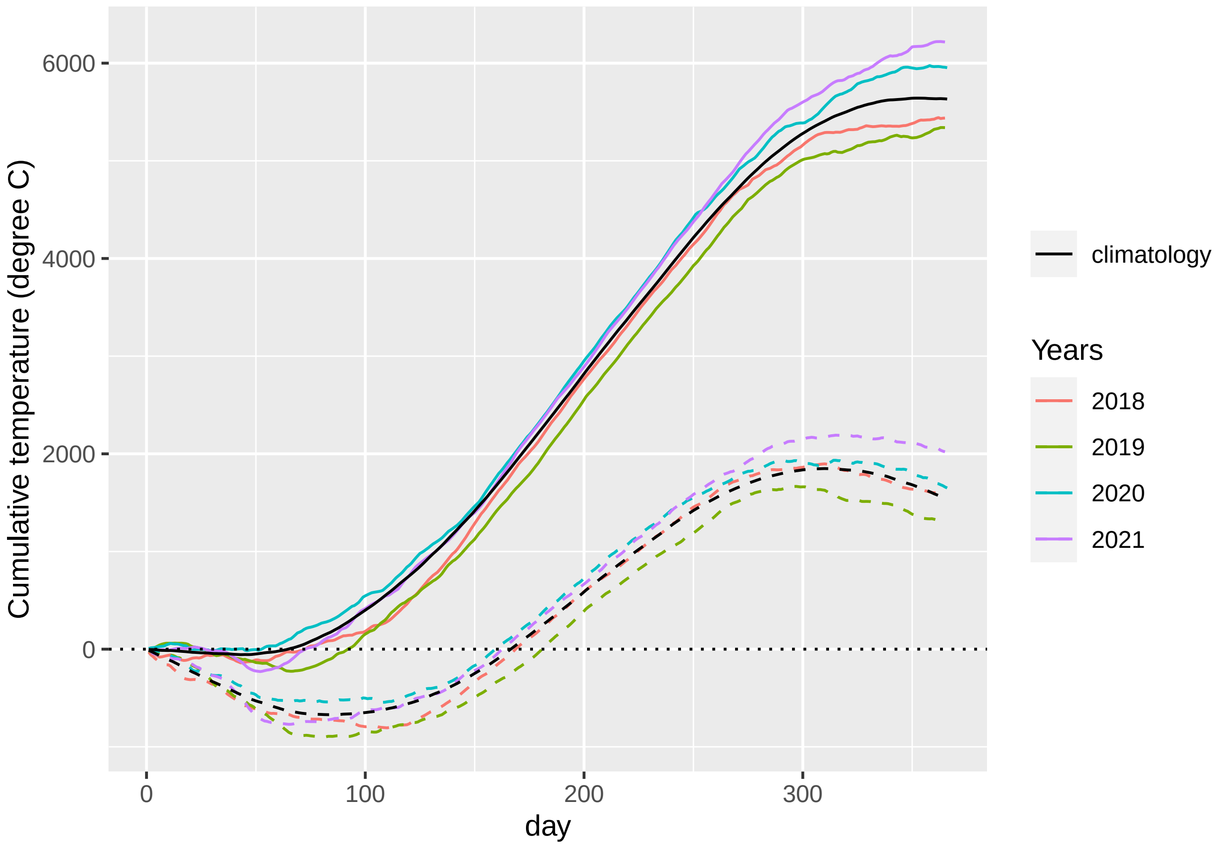

```r
## Cumulative is sometimes easier to see
plot(ames.iem, years = 2012:2015, cumulative = TRUE, climatology = TRUE)
```



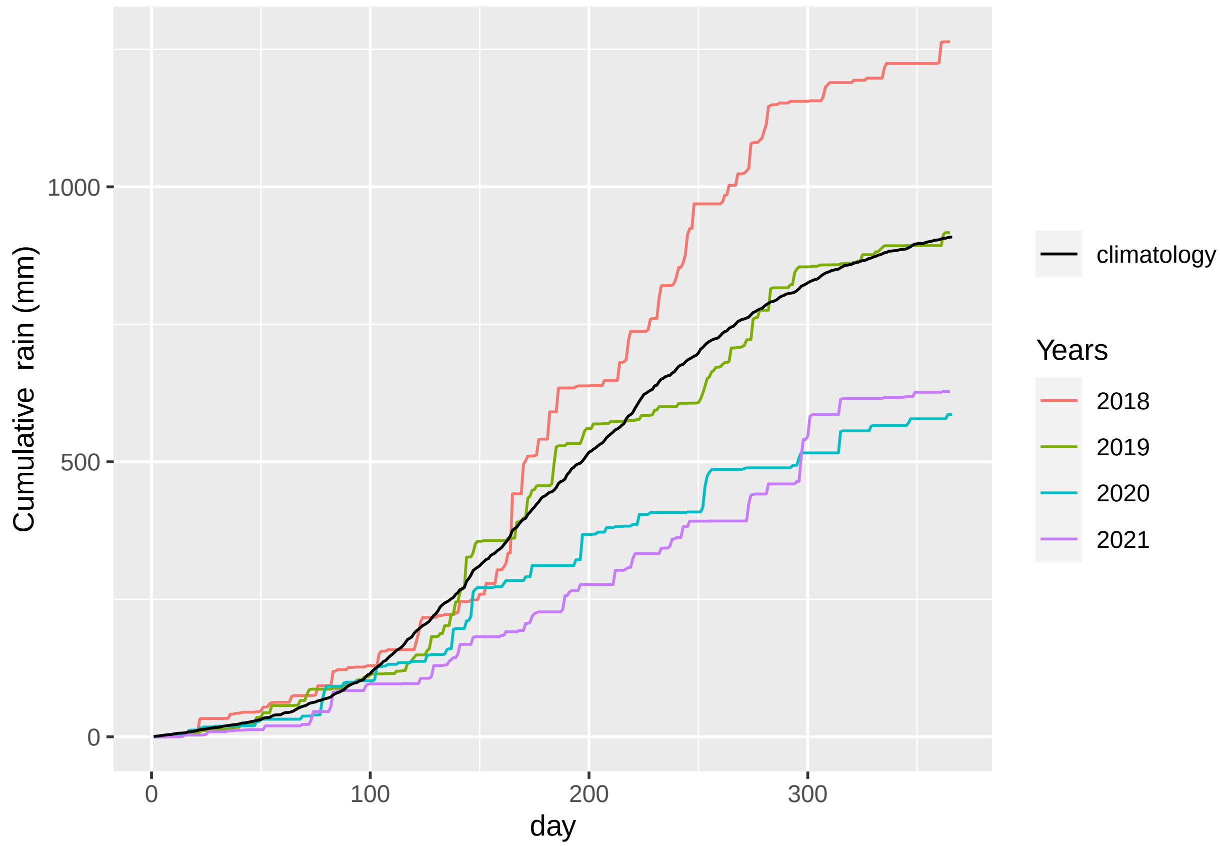

```r
## 2012 stands out as a fairly warm year
## We can also add the climatology
plot(ames.iem,
     met.var = "rain",
     years = 2012:2015,
     cumulative = TRUE,
     climatology = TRUE)
```



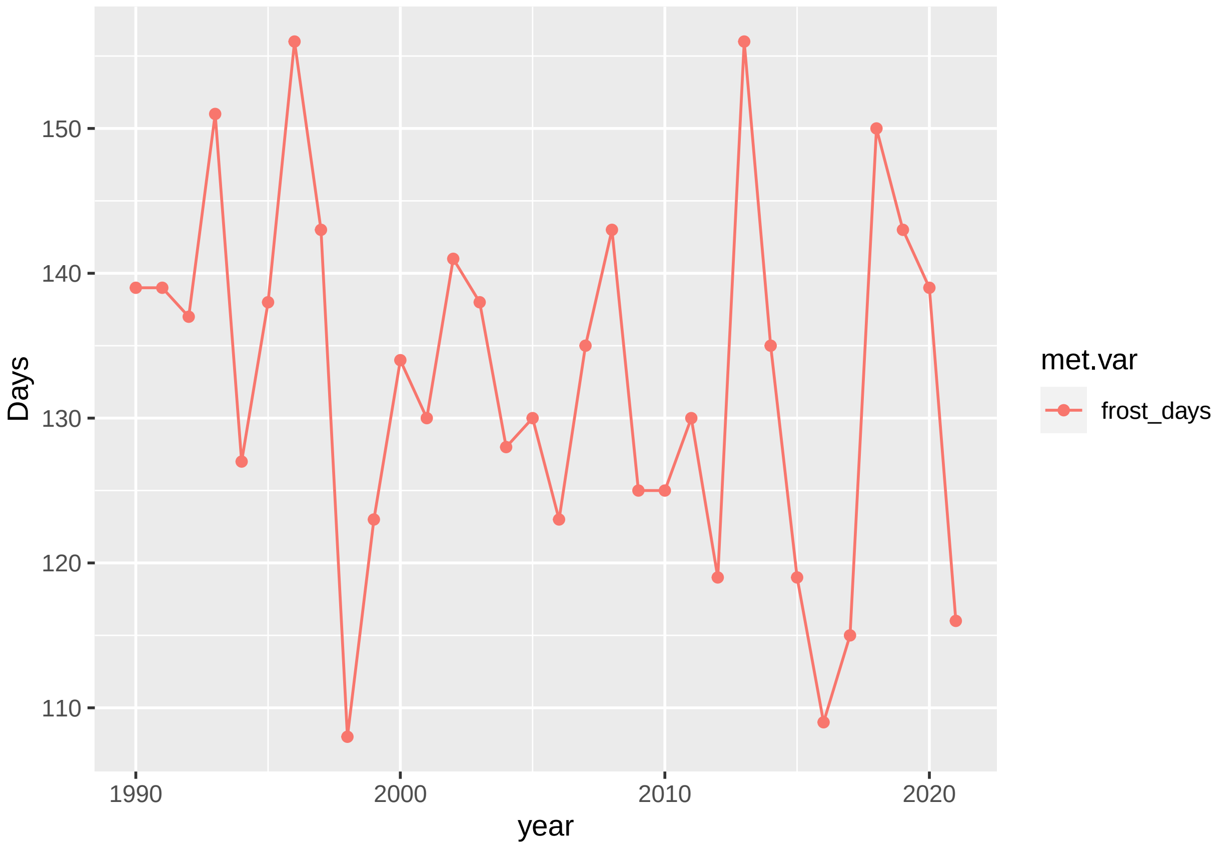

```r
## So clearly 2012 stands out as a HOT and DRY year
## Let's look at more recent data
plot(ames.iem,
     years = 2018:2021,
     cumulative = TRUE,
     climatology = TRUE)
```




```r
plot(ames.iem,
     met.var = "rain",
     years = 2018:2021,
     cumulative = TRUE,
     climatology = TRUE)
```

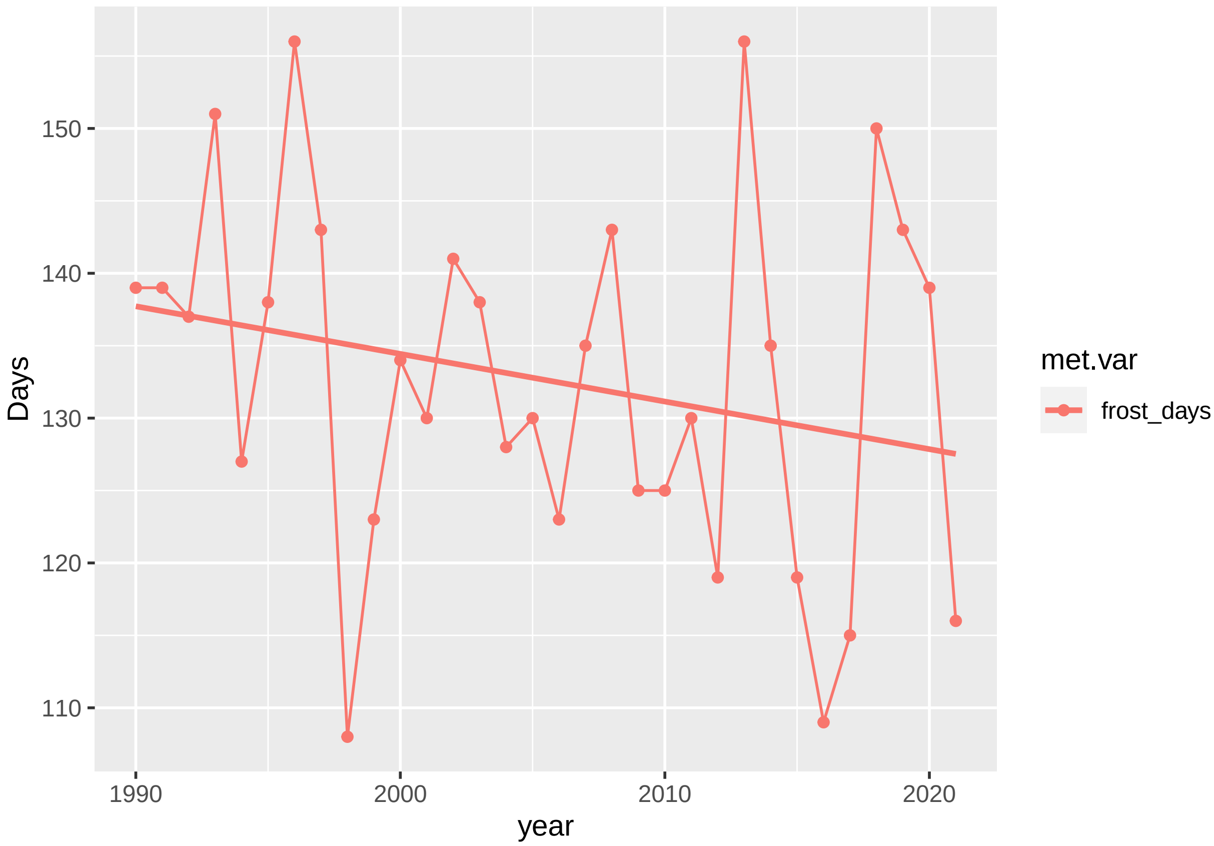



```r
## Combining summary utilities and plotting
## This needs ggplot2
p1 <- plot(ames.iem,
           summary = TRUE,
           compute.frost = TRUE,
           met.var = "frost_days")
```




```
p1 + geom_smooth(method = "lm", se = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



Gridded data

The alternative to station data are gridded products which are typically a combination of observations and models. For point simulation or similar work station data will normally be a better option, as long as it is quality controlled. One popular product can be obtained through NASA-POWER (https://power.larc.nasa.gov/).

In the following exercise we will download data through both the *nasapower* and Iowa Environmental Mesonet (IEM) and perform a comparison and simple visualization.

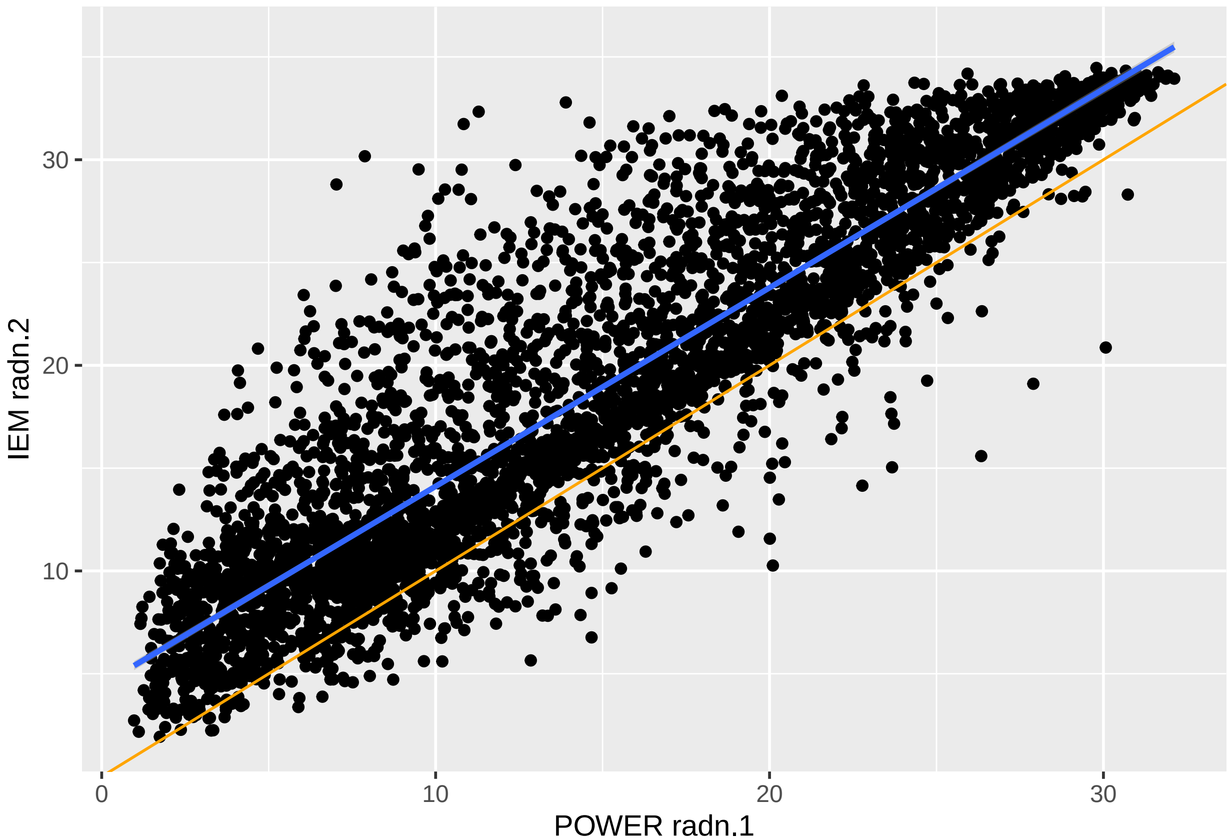
```
library(apsimx)
## Get data close to the energy farm
pwr <- get_power_apsim_met(lonlat = c(-93.77, 42.02),
                           dates = c("2010-01-01","2021-12-10"))

iem <- get_iem_apsim_met(lonlat = c(-93.77, 42.02),
                         dates = c("2010-01-01","2021-12-10"))
```

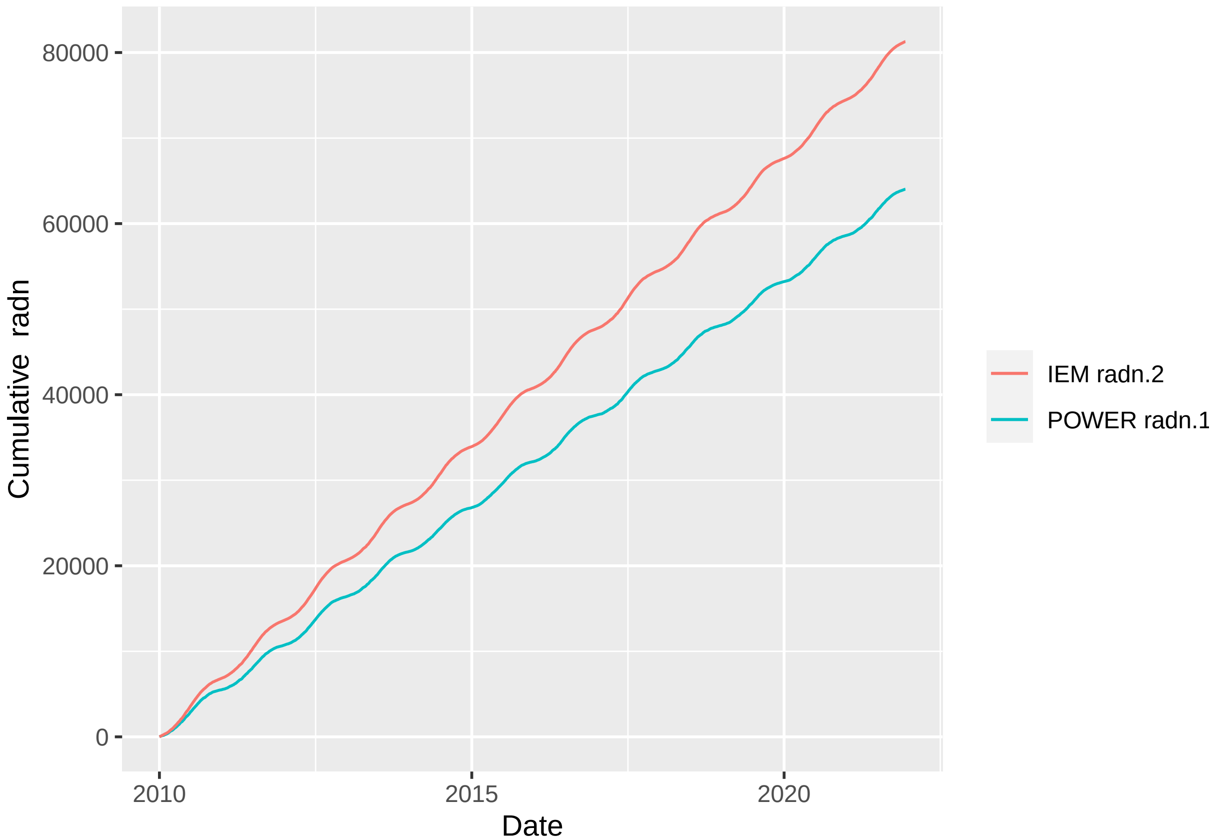
We can compare the solar radiation from both sources.

```
## Comparing variables. We only select the first 6 columns from POWER
cmp <- compare_apsim_met(pwr[, 1:6], iem, labels = c("POWER", "IEM"))
## Let's compare solar radiation
plot(cmp, met.var = "radn") ## IEM has a poitive bias
```

```
## `geom_smooth()` using formula 'y ~ x'
```

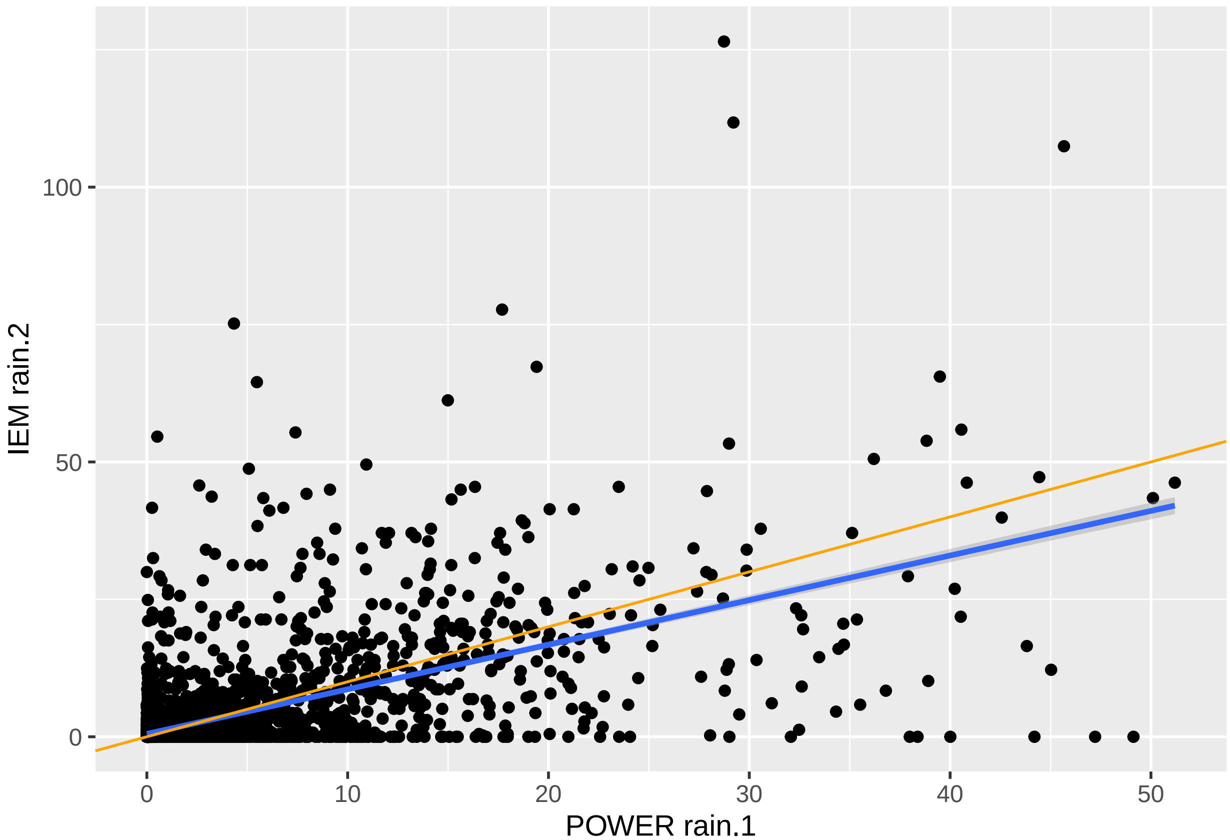



```
plot(cmp, met.var = "radn", plot.type = "ts", cumulative = TRUE)
```

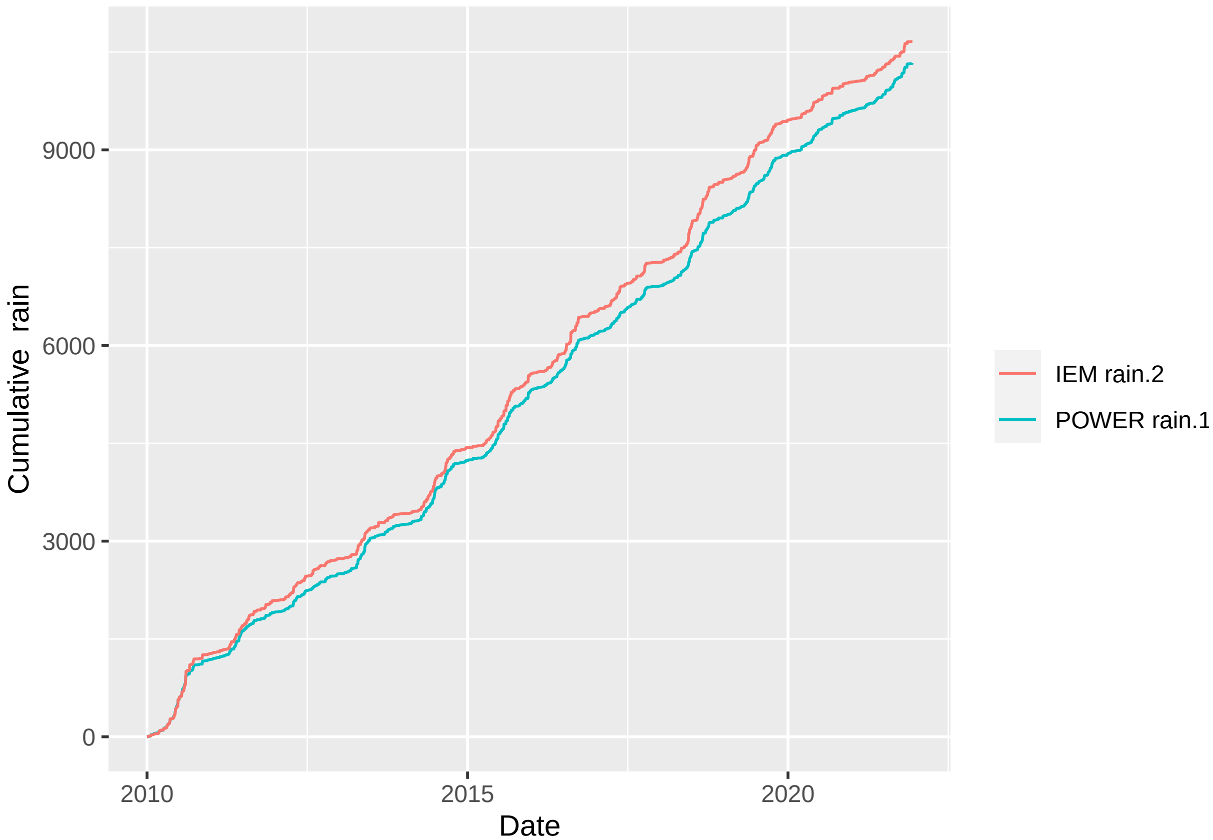



```
## Let's compare precipitation
plot(cmp, met.var = "rain") ## The difference is smaller
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
plot(cmp, met.var = "rain", plot.type = "ts", cumulative = TRUE)
```



Soil Data

US soil database

```
## This line gets data from SSURGO, but just the tables
ams.tbls <- get_ssurgo_tables(lonlat = c(-93.77, 42.02))
## Let's see the structure
names(ams.tbls)
```

```
## [1] "mapunit"     "component"   "chorizon"     "mapunit.shp"
```

```
class(ams.tbls)
```

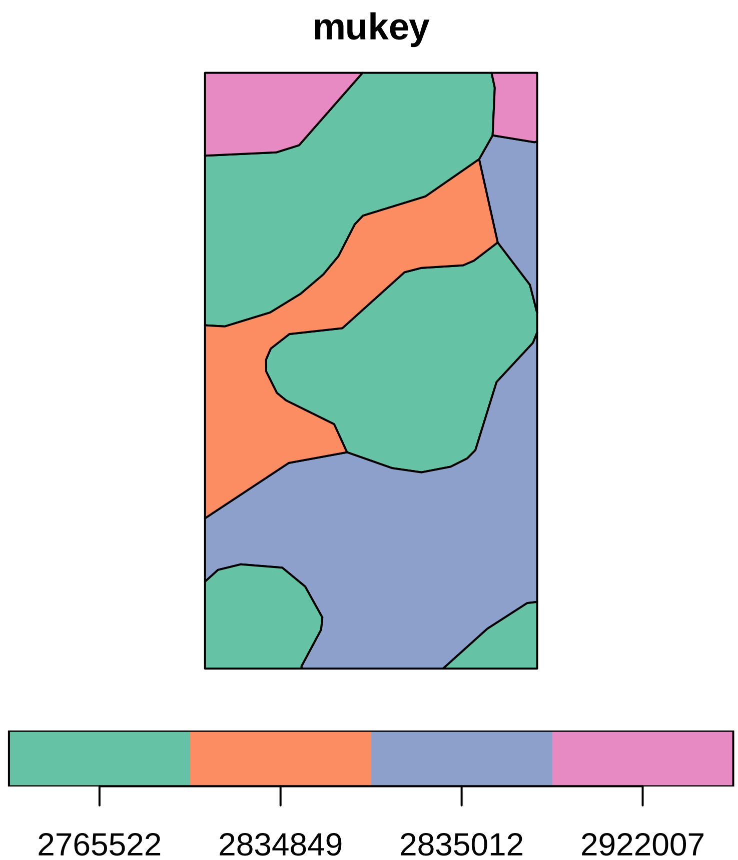
```
## [1] "list"
```

```
ams.tbls$mapunit.shp
```

```
## Simple feature collection with 1 feature and 2 fields
## Geometry type: POINT
## Dimension:     XY
## Bounding box:  xmin: -93.77 ymin: 42.02 xmax: -93.77 ymax: 42.02
## CRS:           +proj=longlat +datum=WGS84 +no_defs
##                geometry   MUKEY AREASYMBOL
## 1 POINT (-93.77 42.02) 2765522      IA015
```

```
## Retrieving an area
ams.tbls2 <- get_ssurgo_tables(lonlat = c(-93.77, 42.02), shift = 300)
```

```r
ams.tbls2$mapunit.shp$mukey <- as.factor(ams.tbls2$mapunit.shp$mukey)
plot(ams.tbls2$mapunit.shp[, "mukey"], key.pos = 1)
```



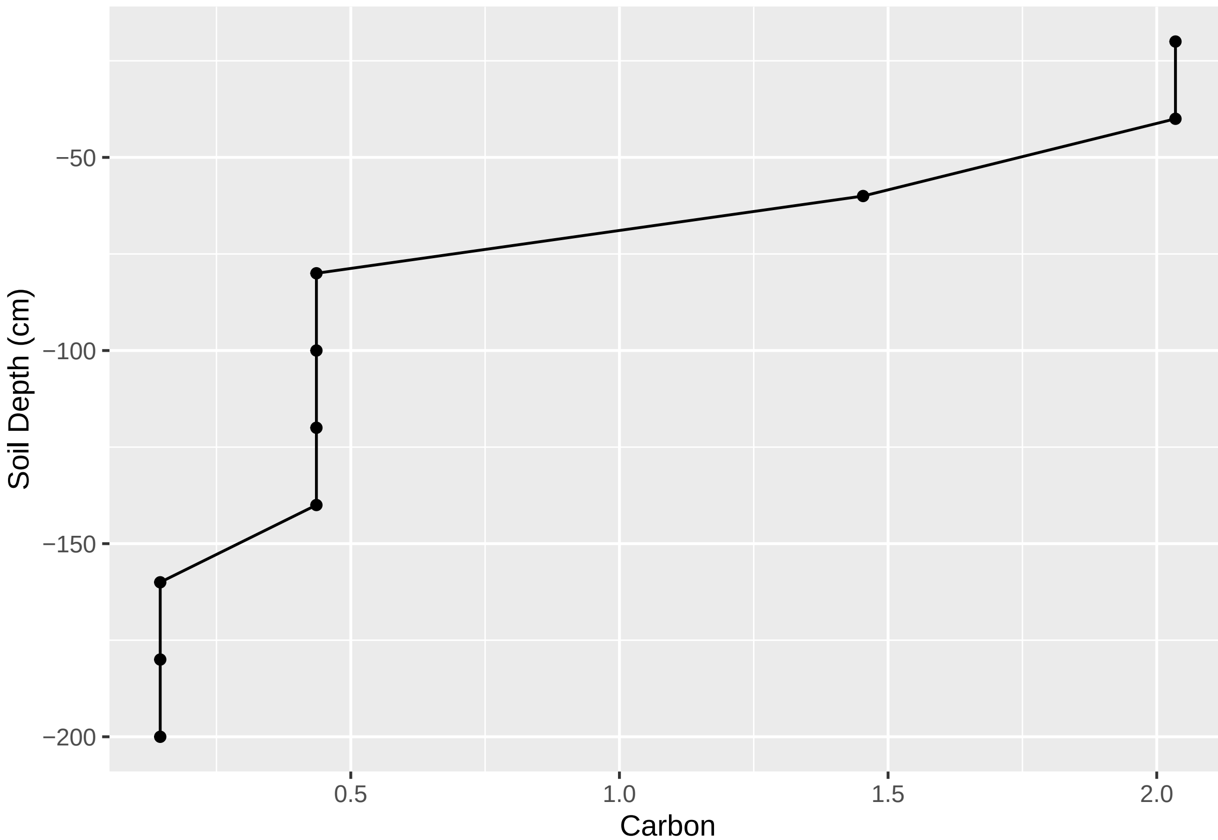

```r
## Looking at soil profiles
sps <- get_ssurgo_soil_profile(lonlat = c(-93.77, 42.02), nsoil = 2)
sps[[1]]$metadata$SoilType
```

```
## [1] "Clarion:2765522"
```

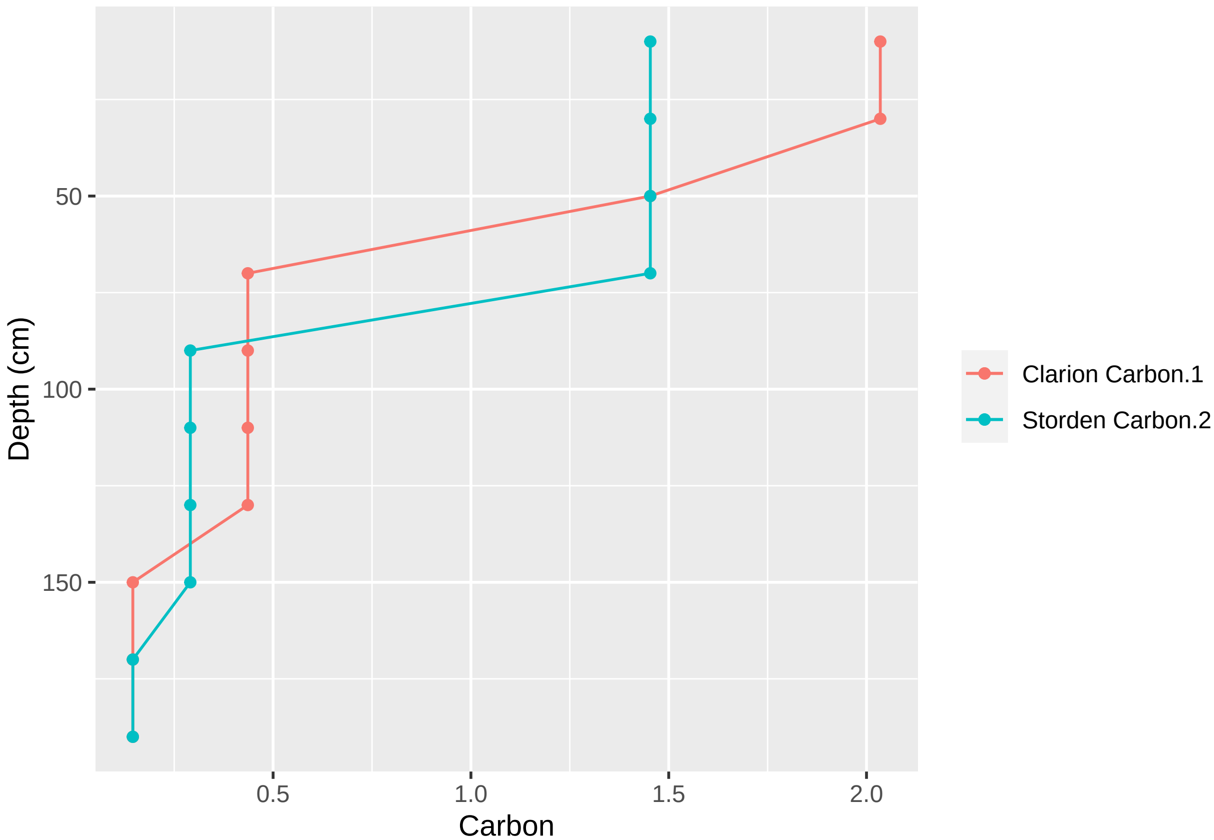
```r
sps[[2]]$metadata$SoilType
```

```
## [1] "Storden:2765522"
```

```r
plot(sps[[1]], property = "Carbon")
```

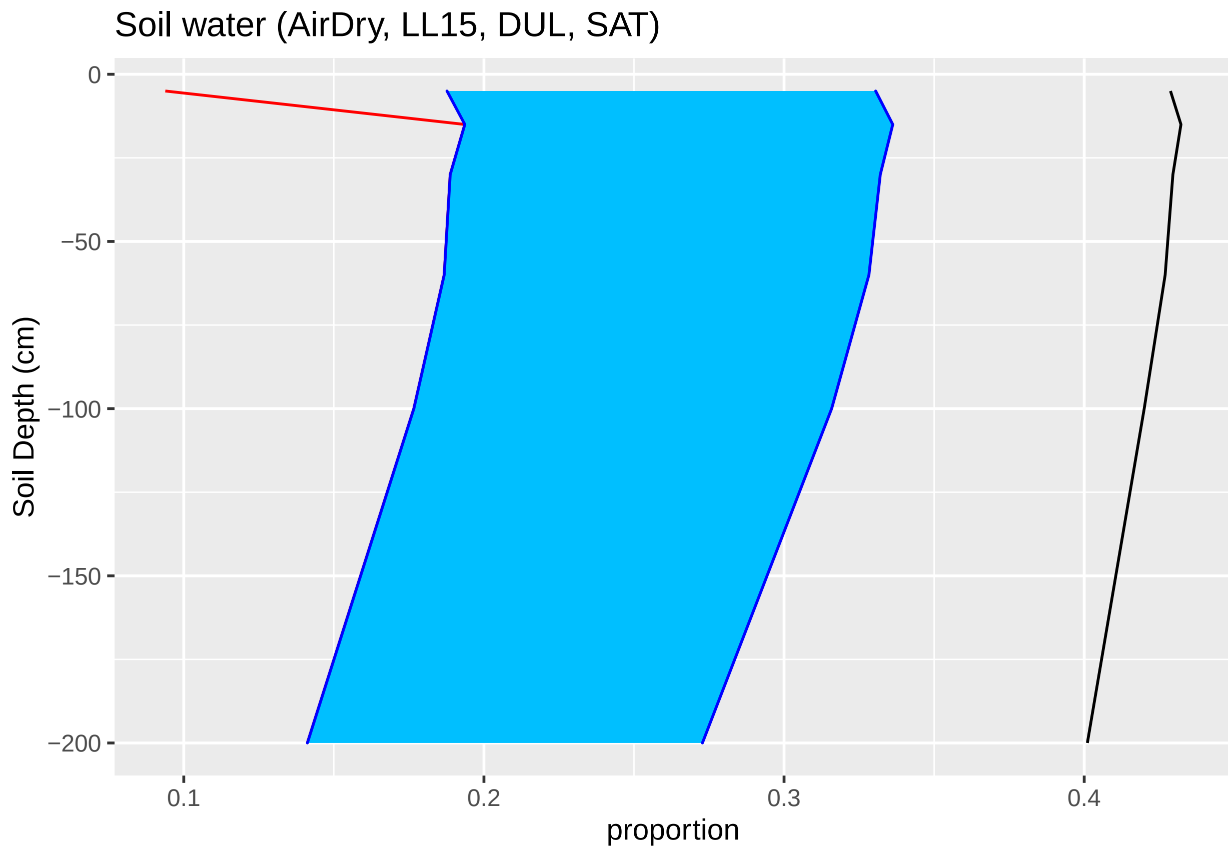



```
cmp.soils <- compare_apsim_soil_profile(sps[[1]],
                                        sps[[2]],
                                        labels = c("Clarion", "Storden"))
plot(cmp.soils, soil.var = "Carbon")
```



SoilGrids (Global Database)

```
ams.sgrds <- get_isric_soil_profile(lonlat = c(-93.77, 42.02))
plot(ams.sgrds, property = "water")
```


Soil water (AirDry, LL15, DUL, SAT)
Soil Depth (cm)
0
−50
−100
−150
−200
0.1
0.2
0.3
0.4
proportion