Weather and Soil Databases for APSIM with R: retrieval, manipulation and visualization

Fernando Miguez, Department of Agronomy, Iowa State University

IOWA STATE UNIVERSITY
Department of Agronomy

APSIM INITIATIVE

Outline

- Sources of weather data
- R activities
- Sources of soil data
- R activities
- Additional Resources

Outline

- Sources of weather data
- R activities
- Sources of soil data
- R activities
- Additional Resources

Weather Data

Station data

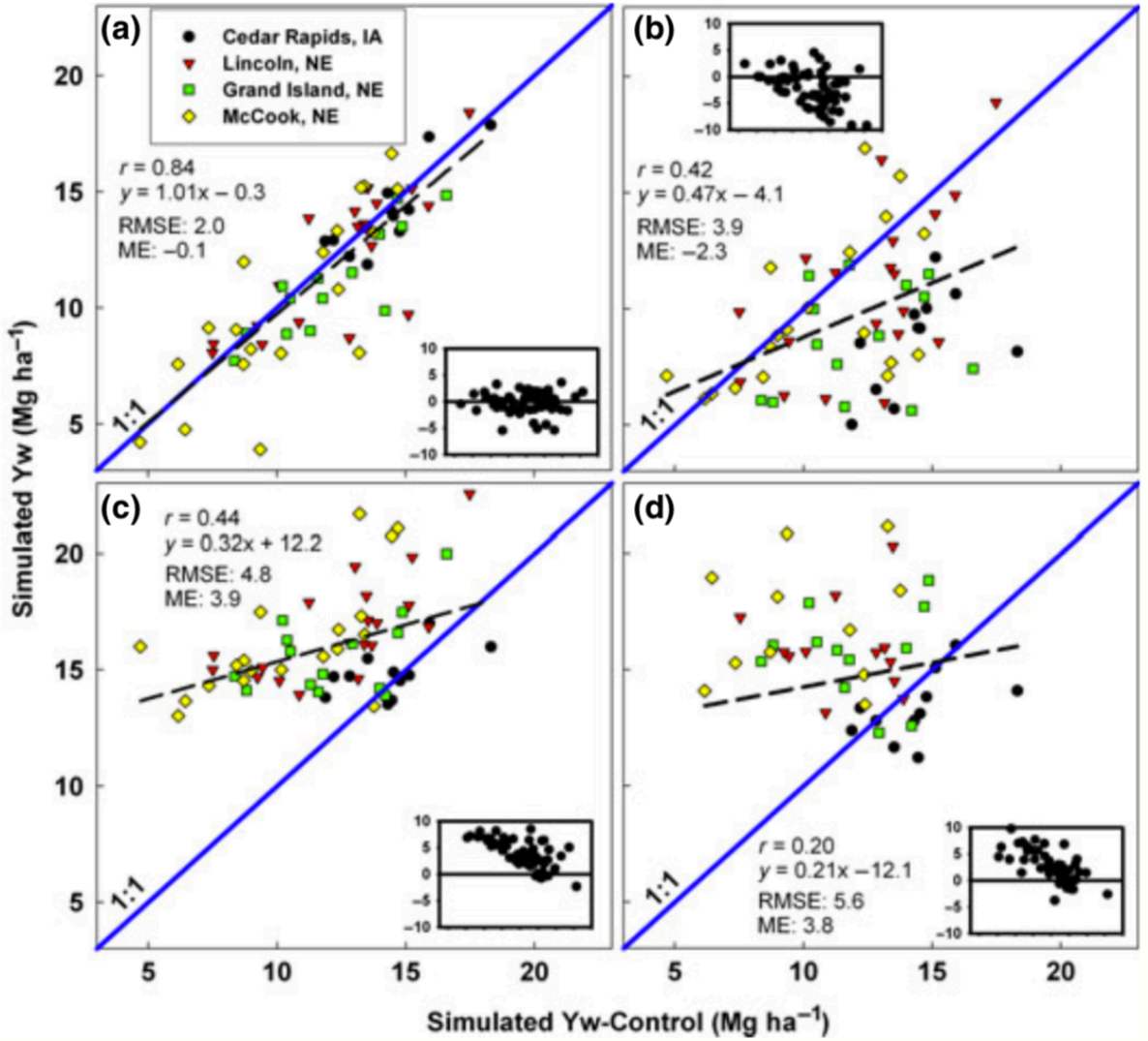
- Better for point simulation
- Need to QC
- Need to consider location
- Often incomplete

Gridded/Modeled data (GWD)

- Better for regional scale
- Grid scale (resolution)
- Temporal resolution
- Bias for point simulations
- Often better source for solar radiation

Impact of derived global weather data on simulated crop yields

JUSTIN VAN WART, PATRICIO GRASSINI and KENNETH G. CASSMAN
Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583-0915, USA



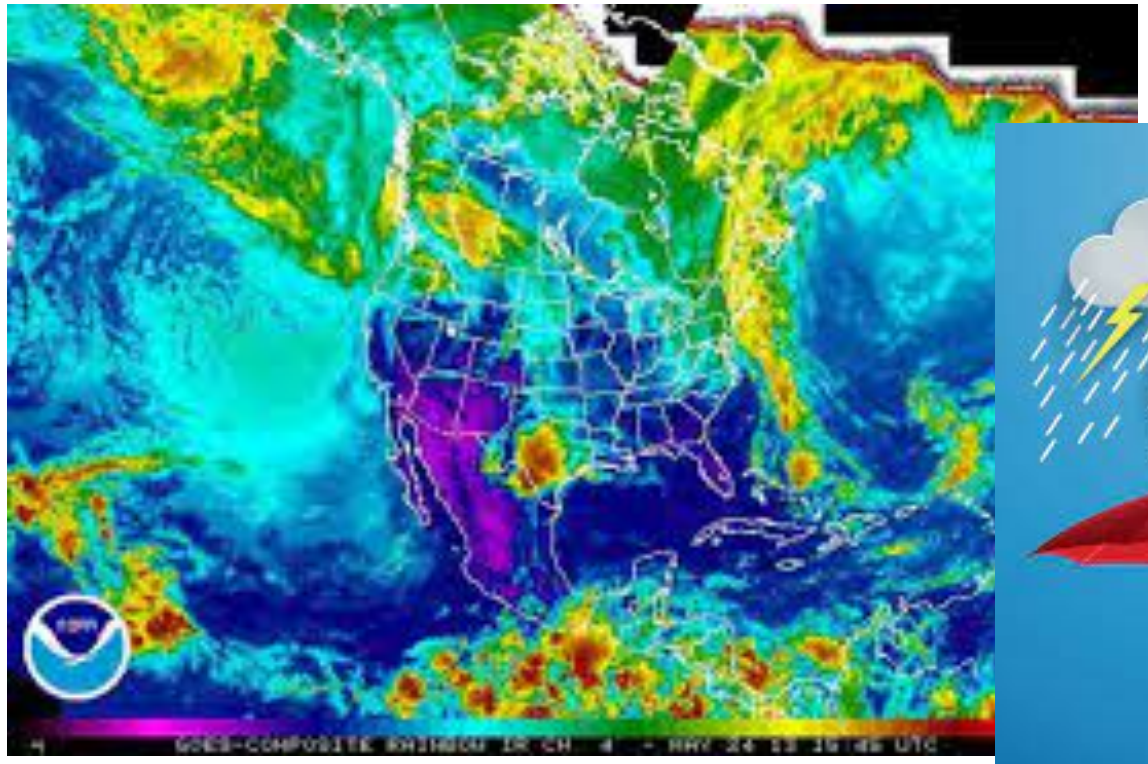
a) NOAA-SR
b) NCEP
c) CRU
d) NASA

Broad recommendations

- When performing point simulations
  - Obtain station data if you can
  - Need for quality control
  - Compare solar radiation data or replace with modeled SR (NASA-POWER)
  - Might need to supplement data from other sources

- When performing regional simulations
  - If possible, run simulations using more than one source
  - Perform an assessment of bias and variance

Weather Data Sources



Iowa Environmental Mesonet (Not just Iowa)

1. Select station(s):

Select Widget for IACLIMATE Network

ℹ Station Metadata

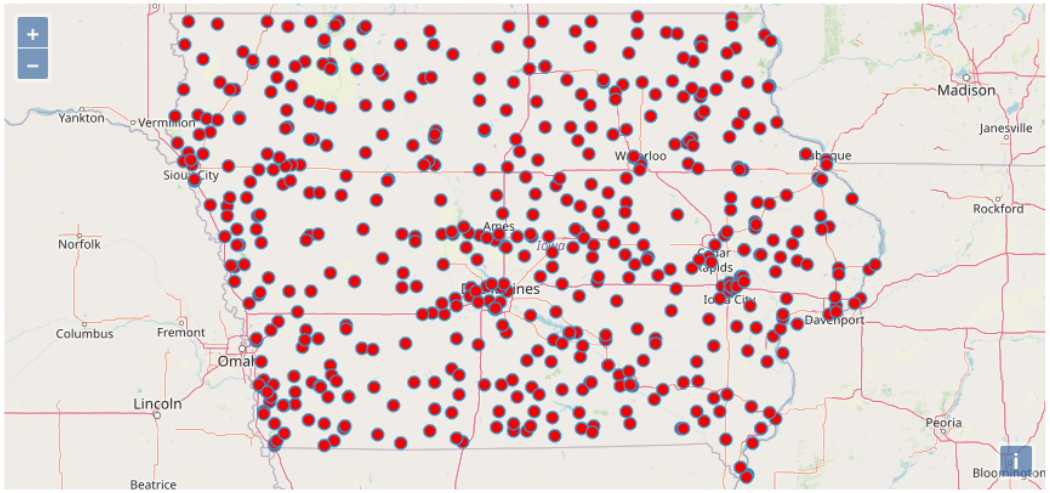
Sort Available Stations:

[IA0000] Iowa Average (1893-Now)
[IA0064] AFTON (1894-1950)
[IA0070] ADEL 1 SE (1958-1967)
[IA0088] AKRON (1900-Now)

Enter some text here to | + Add Selected | Add All

Selected Stations:

− Remove Selected | Remove All



2. Select Start/End Time:

Please note the start year in the station selector. If you only want data for one day, set the start and end times to the same value.

	Year	Month	Day
Start:	1893	January	1
End:	2021	August	26

3. Select Variables / Model Input Format:

Select either one or more data variables

Computed: Growing Degree Days (base=50,max=86) [F]
Computed: Growing Degree Days (base=40,max=86) [F]
Variable: High Temperature [F]
Variable: High Temperature [C]
Variable: Low Temperature [F]

4. Specialized Data/Model Formats

When you select a given data format, it will override any selected variables above and formats below. These are specialized formats typically used for modelling and their choice dictates the variables to be included. Please contact us to have your format added to the list!

Model: APSIM MET File
Model: Century (Monthly Format)
Model: DailyDayCent (no extra drivers)
Model: DNDC
Model: SALUS

5. How to view?

Excel File

6. Data Delimitation:

This option is only relevant if you did not select "Excel File" from #5 above.

Comma

7. Include Latitude and Longitude in output?

No

8. Supplement 2021's data with scenario year

https://mesonet.agron.iastate.edu/request/coop/fe.phtml

Data sources

- IEM (https://mesonet.agron.iastate.edu/) – Station and Gridded (US)
- NASA POWER (https://power.larc.nasa.gov/) – Gridded (Global)
- DayMet (https://daymet.ornl.gov/) – Gridded (US)
- NOAA-GSOD (https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00516) – Station and Model (Global)

Weather data with R

- R (4.0.0 or higher)
- apsimx R package (version 2.3.1)

- This package relies on other packages for weather
  - **nasapower**
  - daymetr
  - GSODR
  - chirps
  - FedData

Outline

- Sources of weather data
- R activities
  - Retrieving
  - Summarizing
  - Visualizing
  - Comparing
- Sources of soil data
- R activities
- Additional Resources

```
library(apsimx)
library(ggplot2)
```

Weather data with R

```
## 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)
## Quick visualization for just a few years, but it is still hard to see
plot(ames.iem, years = 2012:2015)
## Cumulative is sometimes easier to see
plot(ames.iem, years = 2012:2015, cumulative = TRUE, climatology = TRUE)
```

Outline

- Sources of weather data
- R activities
  - Retrieving
  - Summarizing
  - Visualizing
  - Comparing
- Sources of soil data
- R activities
- Additional Resources

Summary for object of class ‘met’

```
> summary(ames.iem, years = 2018:2021)
  year months days high_maxt high_mint avg_maxt avg_mint low_maxt low_mint rain_sum radn_sum radn_avg
1 2018   1:12 1:31      36.1      23.3    14.90     4.29    -21.7    -30.0  1264.17  6727.35    18.43
2 2019   1:12 1:31      34.4      23.9    14.63     3.60    -20.6    -30.6   916.69  6365.31    17.44
3 2020   1:12 1:31      34.4      24.4    16.27     4.51    -16.1    -24.4   585.75  6880.47    18.80
4 2021   1:12 1:31      38.3      25.0    17.07     5.56    -21.7    -32.2   627.09  6964.74    19.08
> ?summary.met
```

Summary for an APSIM met file

Description

Create a data.frame summarizing an object of class ‘met’

Usage

```
## S3 method for class 'met'
summary(
  object,
  ...,
  years,
  months,
  days,
  julian.days,
  compute.frost = FALSE,
  frost.temperature = 0,
  check = FALSE,
  verbose = FALSE,
  na.rm = FALSE,
  digits = 2
)
```

Arguments

`object`	object of class ‘met’
`...`	optional argument (none used at the momemt)
`years`	optional argument to subset years
`months`	optional argument to subset by months. If an integer, it should be between 1 and 12. If a character, it can be in the format, for example, ‘jan’ or ‘Jan’.
`days`	optional argument to subset by days. It should be an integer between 1 and 31.
`julian.days`	optional argument to subset by julian days. It should be a vector of integers between 1 and 365. Either use ‘days’ or ‘julian.days’ but not both.
`compute.frost`	logical (default FALSE). Whether to compute frost statistics.
`frost.temperature`	value to use for the calculation of the frost period (default is zero).
`check`	logical (default FALSE). Whether to ‘check’ the ‘met’ object.
`verbose`	whether to print additional infomation to the console
`na.rm`	whether to remove missing values. Passed to ‘aggregate’
`digits`	digits for rounding (default is 2).

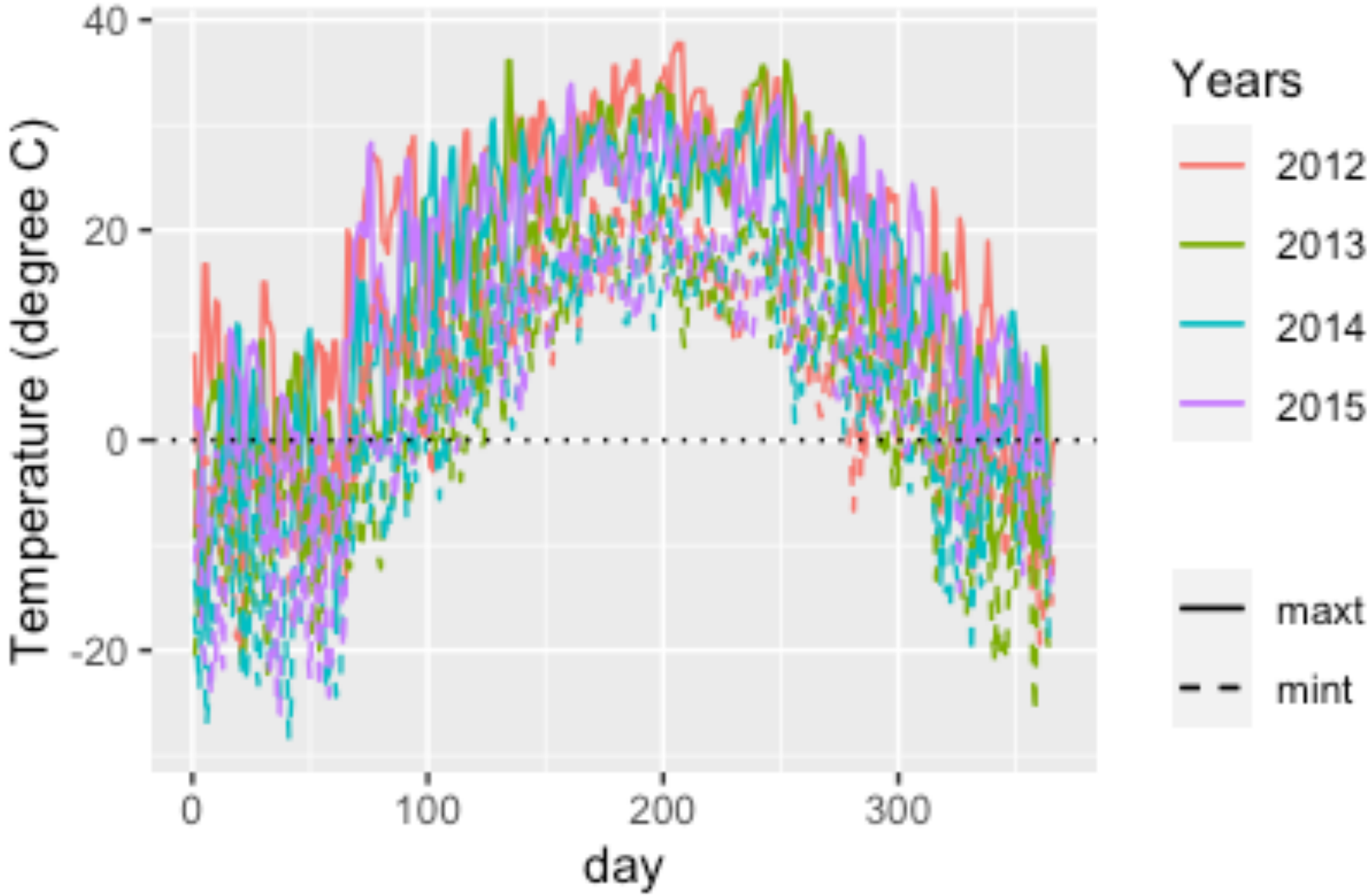
Outline

- Sources of weather data
- R activities
  - Retrieving
  - Summarizing
  - Visualizing
  - Comparing
- Sources of soil data
- R activities
- Additional Resources

```
library(apsimx)
library(ggplot2)
```

Weather data with R

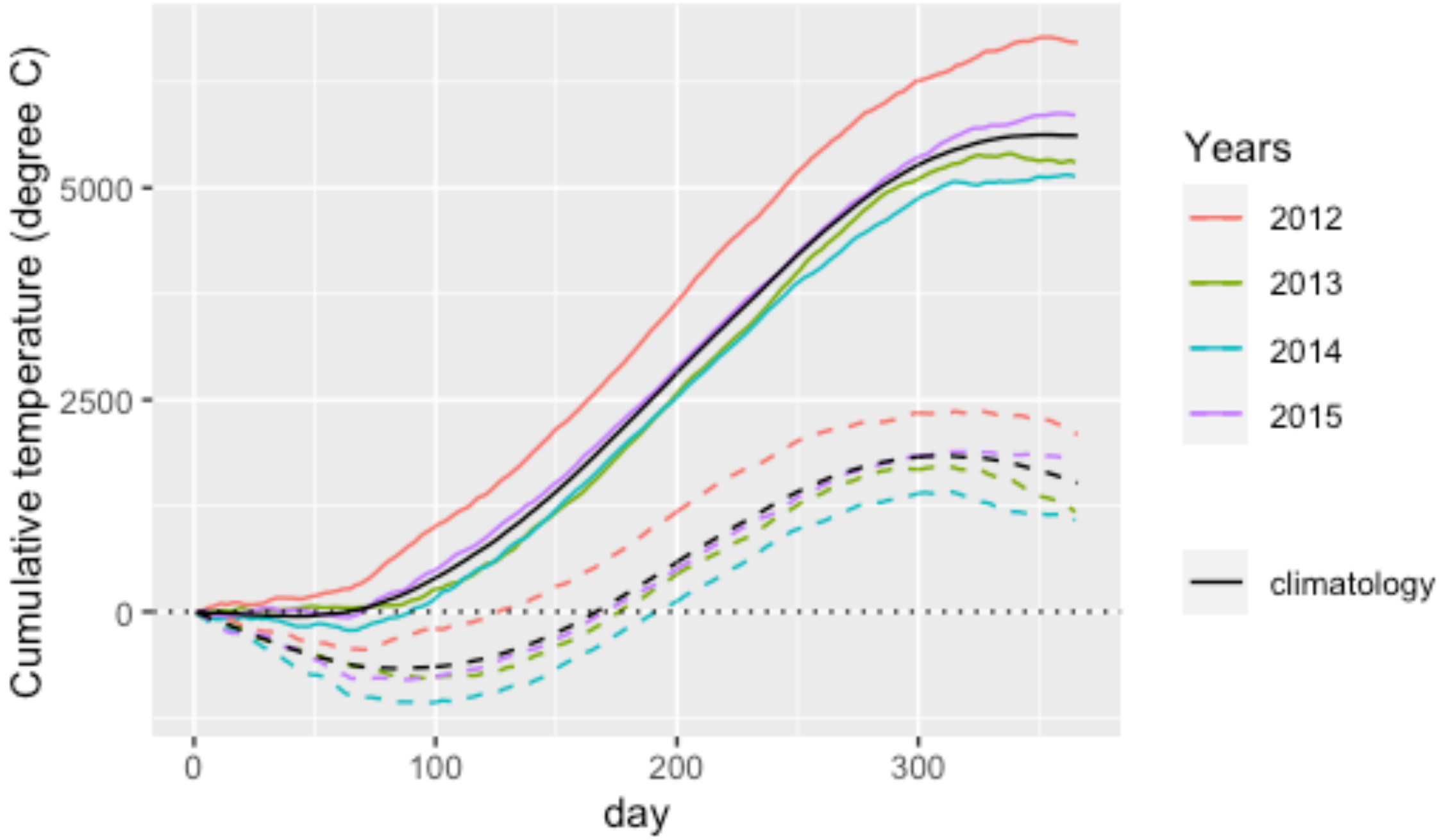
```
## This code 'gets' the data through a call to IEM
ames.iem <- get_iem_apsim_met(lonlat = c(-93.77, 42.02),
                              dates = c("1990-01-01", "2021-12-31"))
## There is a summary function which provides some summaries per year
summary(ames.iem)
## Quick visualization for just a few years, but it is still hard to see
plot(ames.iem, years = 2012:2015)
## Cumulative is sometimes easier to see
plot(ames.iem, years = 2012:2015, cumulative = TRUE, climatology = TRUE)
```


40
20
0
-20
Temperature (degree C)
0
100
200
300
day
Years
2012
2013
2014
2015
maxt
mint

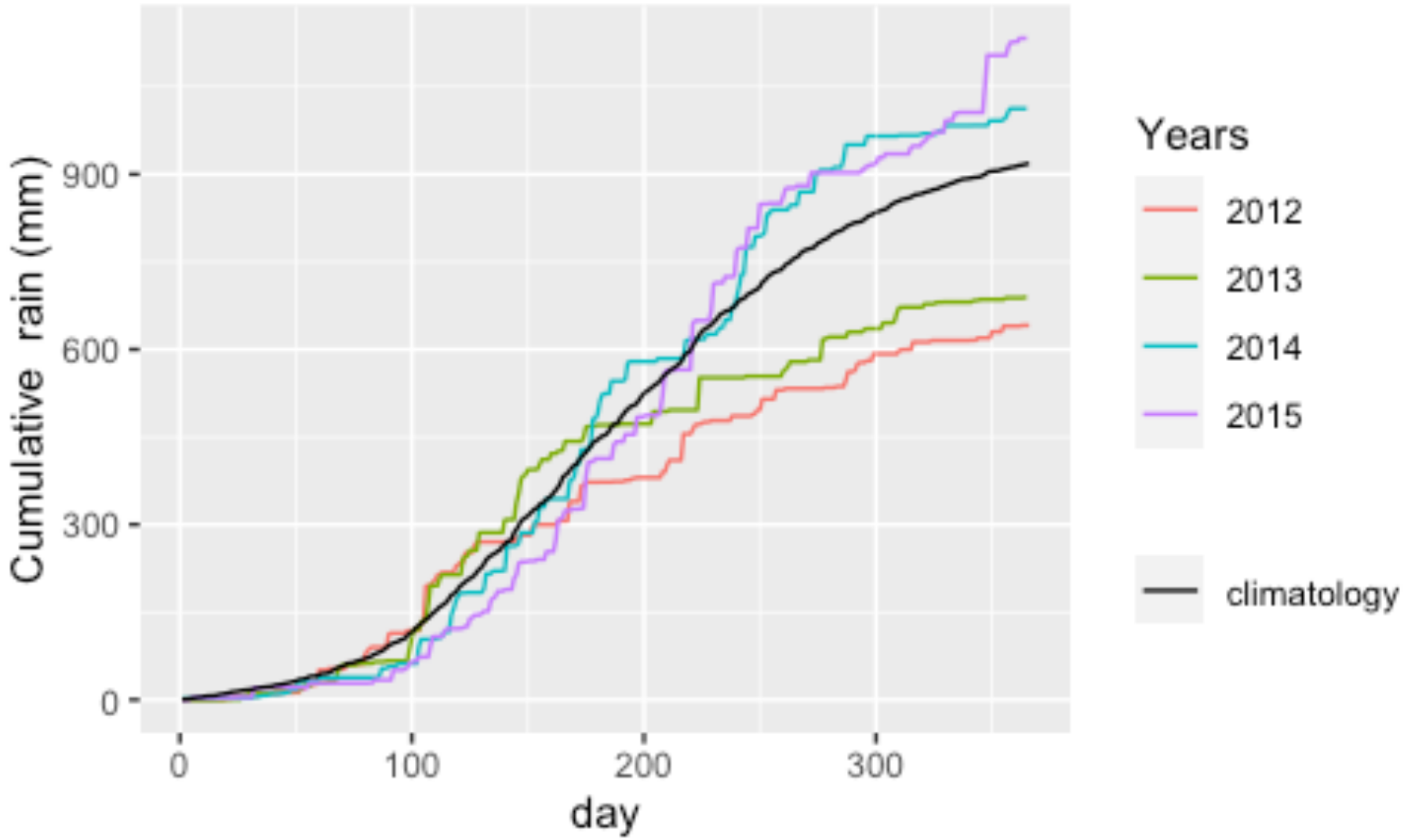
```
library(apsimx)
library(ggplot2)
```

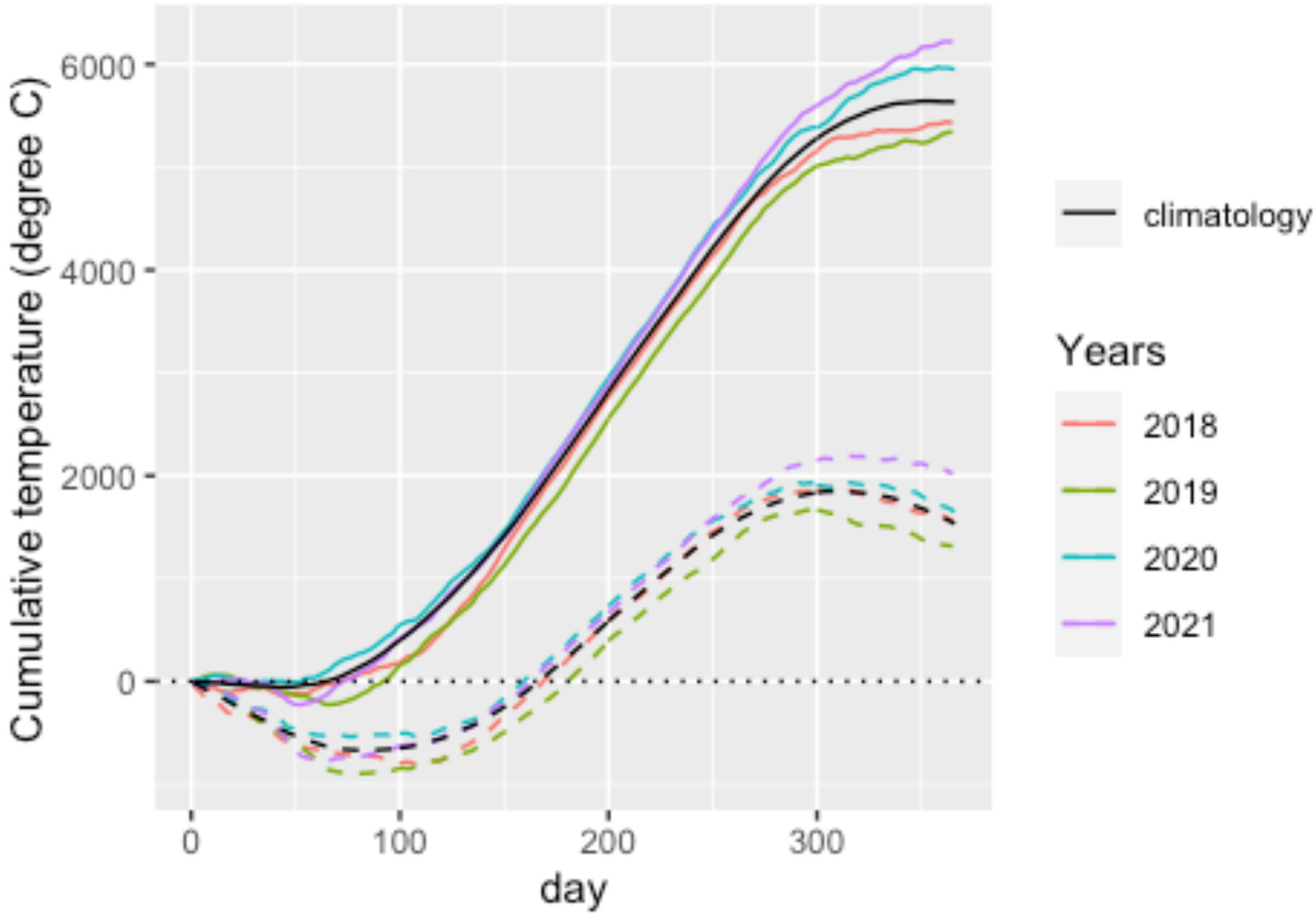
Weather data with R

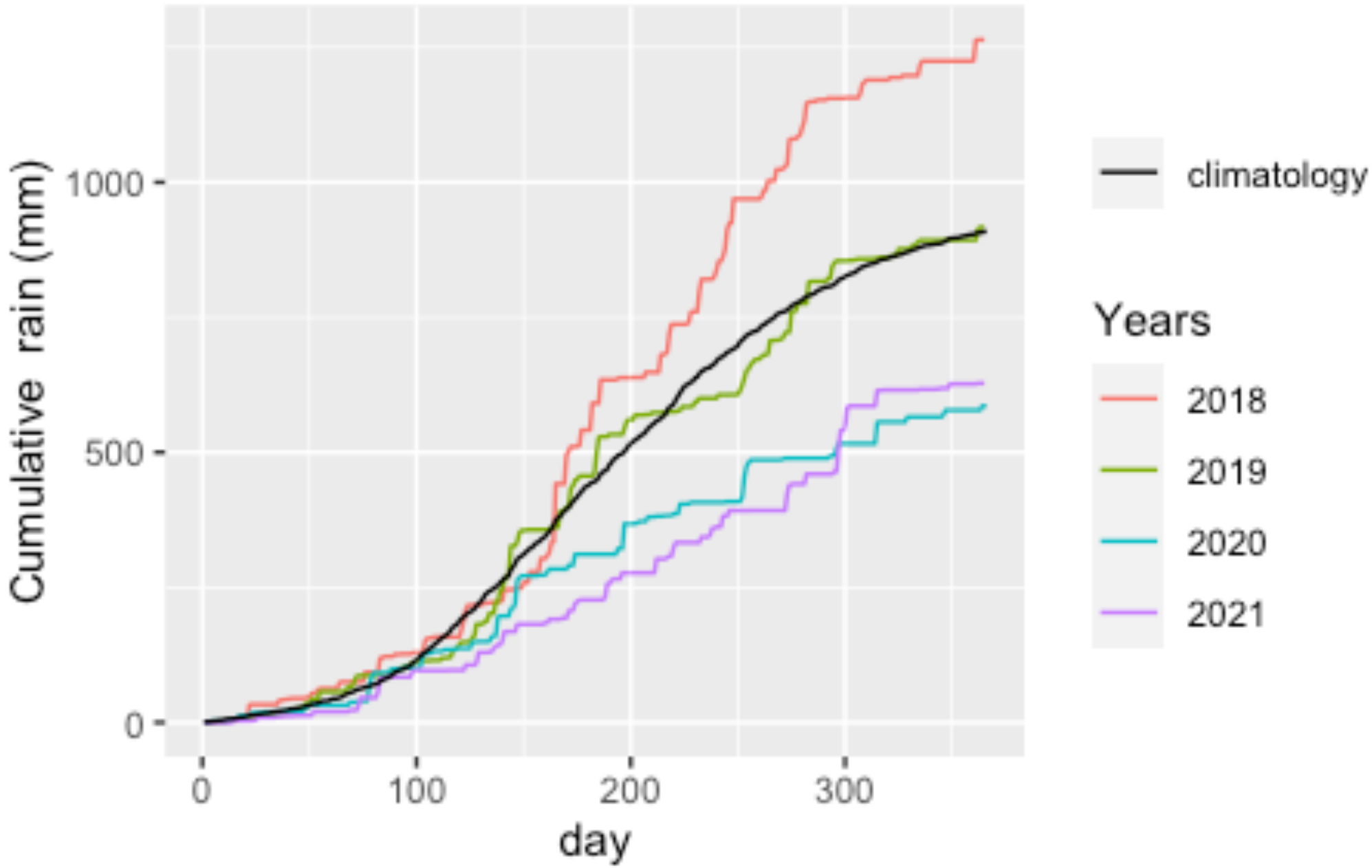
```
## This code 'gets' the data through a call to IEM
ames.iem <- get_iem_apsim_met(lonlat = c(-93.77, 42.02),
                              dates = c("1990-01-01", "2021-12-31"))
## There is a summary function which provides some summaries per year
summary(ames.iem)
## Quick visualization for just a few years, but it is still hard to see
plot(ames.iem, years = 2012:2015)
## Cumulative is sometimes easier to see
plot(ames.iem, years = 2012:2015, cumulative = TRUE, climatology = TRUE)
```


Cumulative temperature (degree C)
5000
2500
0
0
100
200
300
day
Years
2012
2013
2014
2015
climatology

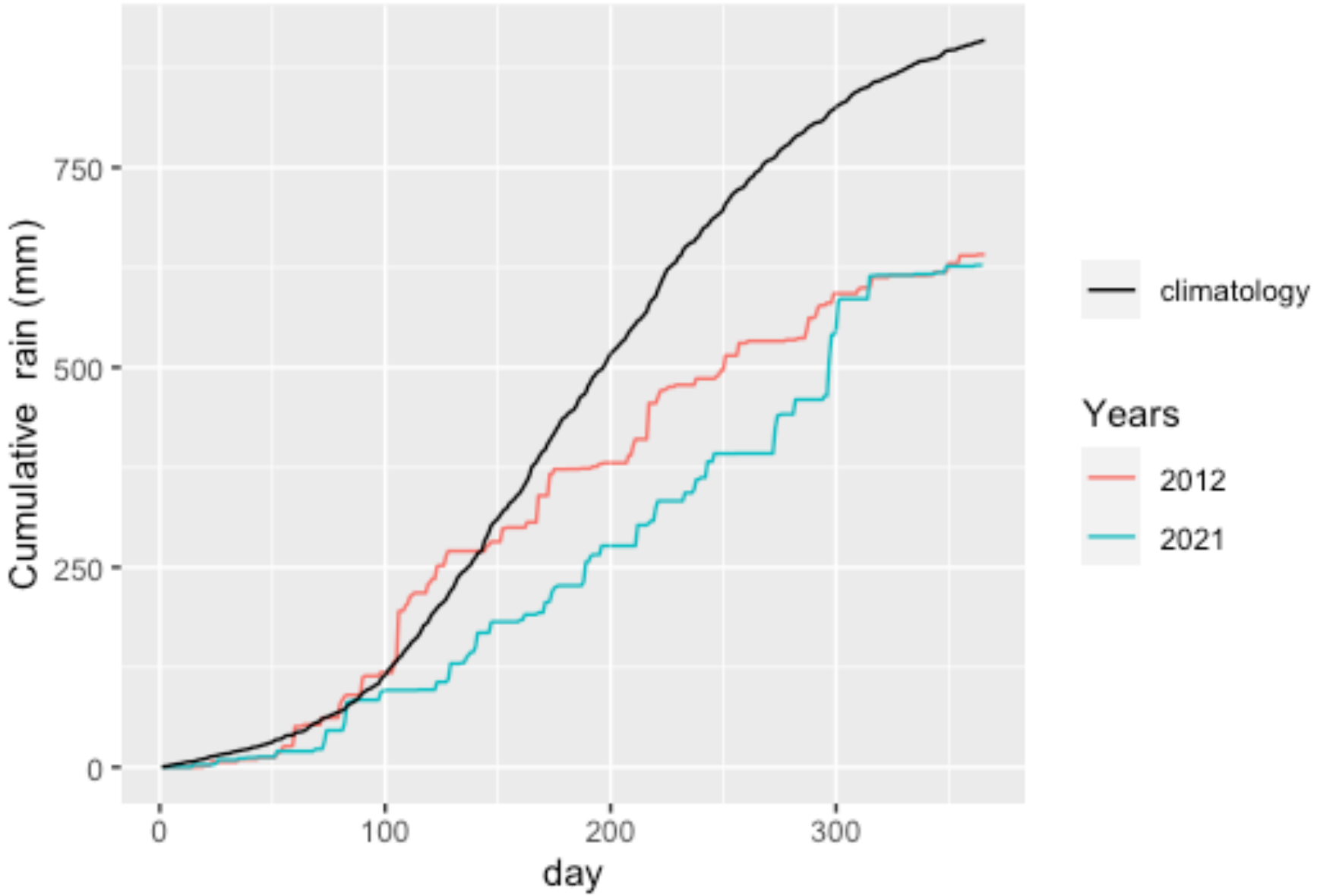
```
plot(ames.iem,
     met.var = "rain",
     years = 2012:2015,
     cumulative = TRUE,
     climatology = TRUE)
```




Cumulative temperature (degree C)
6000
4000
2000
0
0
100
200
300
day
climatology
Years
2018
2019
2020
2021

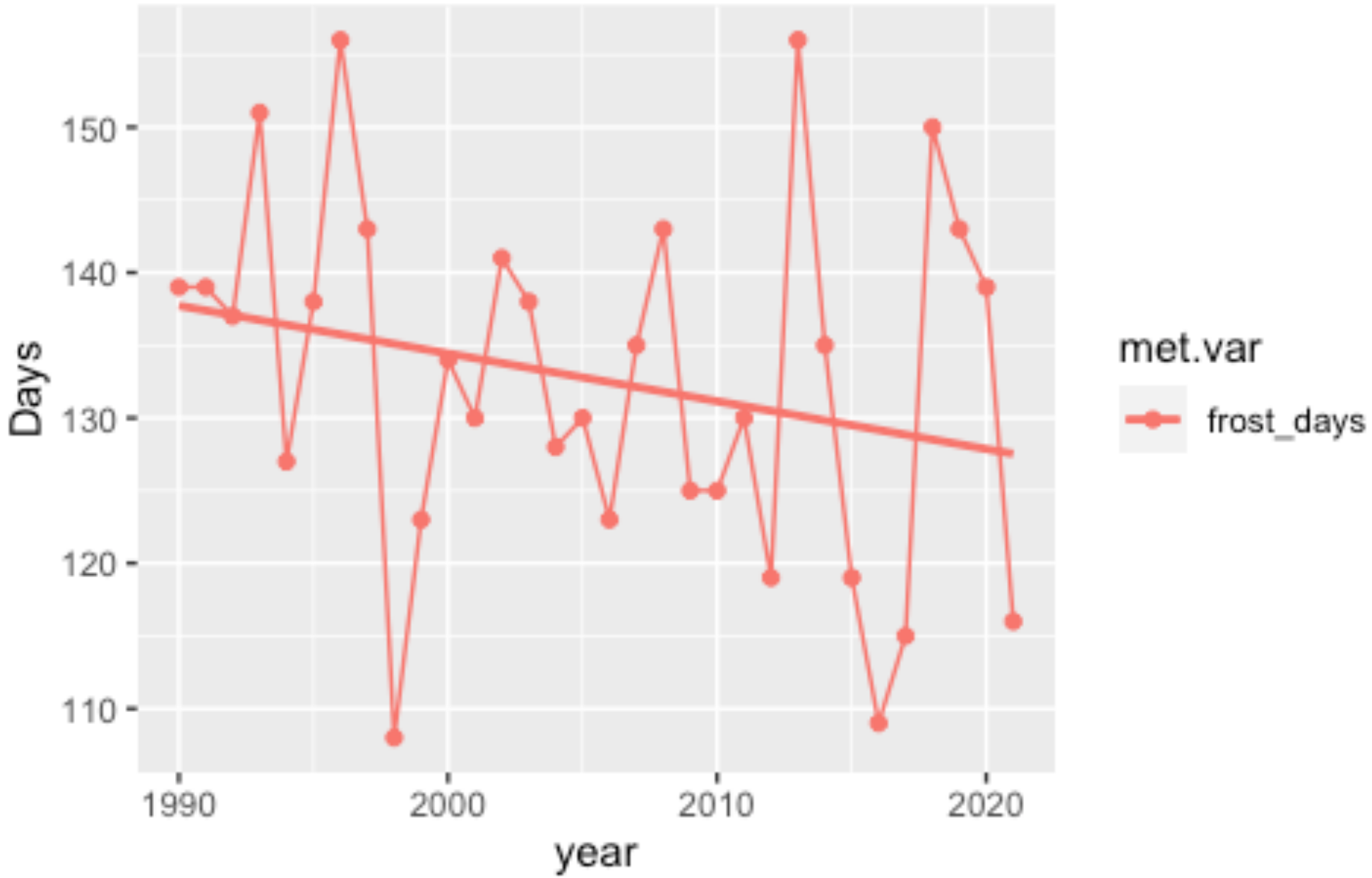

Cumulative rain (mm)
1000
500
0
0
100
200
300
day
climatology
Years
2018
2019
2020
2021

```r
plot(ames.iem,
     met.var = "rain",
     years = c(2012,2021),
     cumulative = TRUE,
     climatology = TRUE)
```



ggplot2 compatible

```
## Combining summary utilities and plotting
library(ggplot2)
p1 <- plot(ames.iem,
           summary = TRUE,
           compute.frost = TRUE,
           met.var = "frost_days")
p1 + geom_smooth(method = "lm", se = FALSE)
```


150
140
130
120
110
Days
1990
2000
2010
2020
year
met.var
frost_days

Outline

- Sources of weather data
- R activities
  - Retrieving
  - Summarizing
  - Visualizing
  - Comparing
- Sources of soil data
- R activities
- Additional Resources

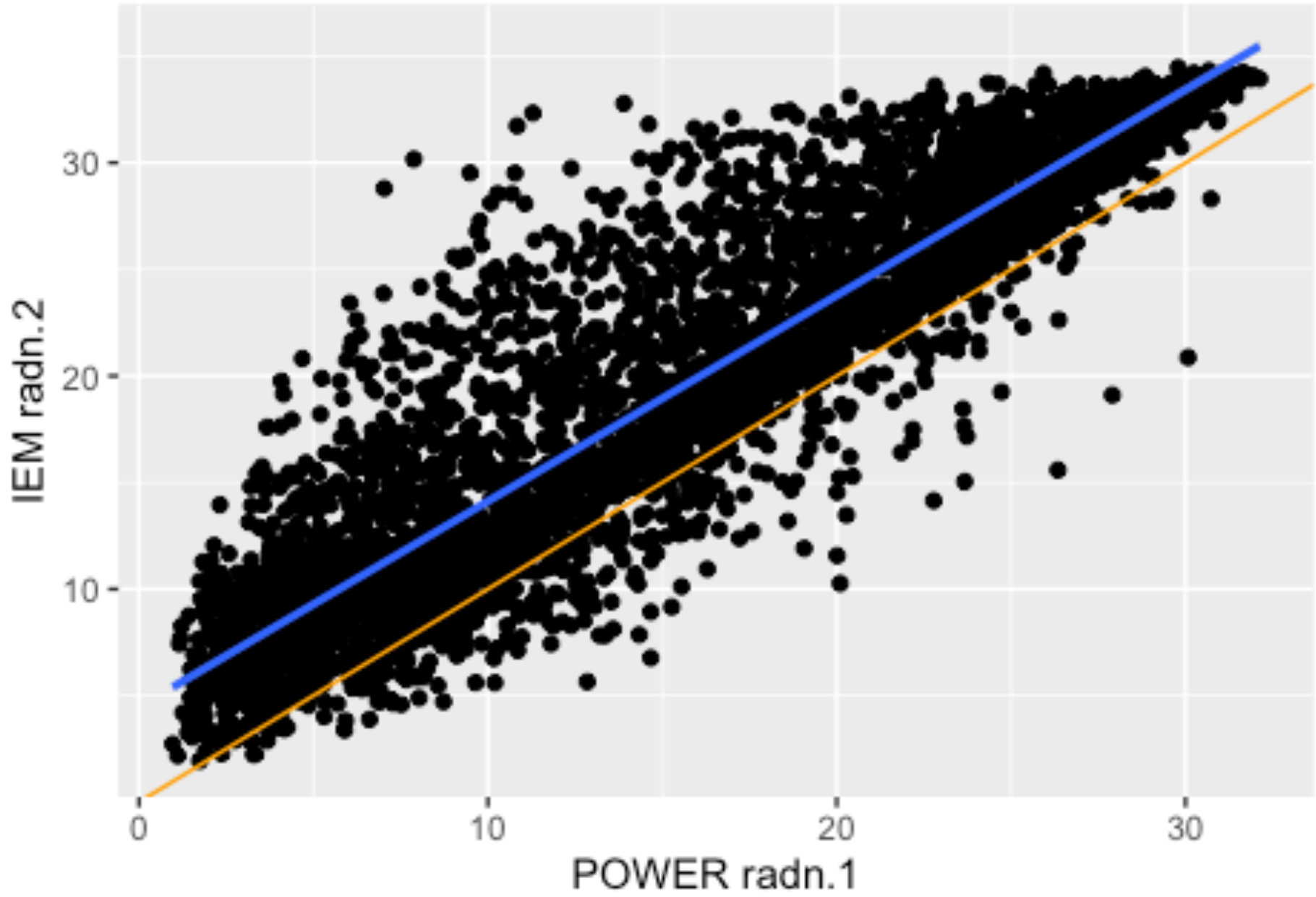
Comparing data from different sources

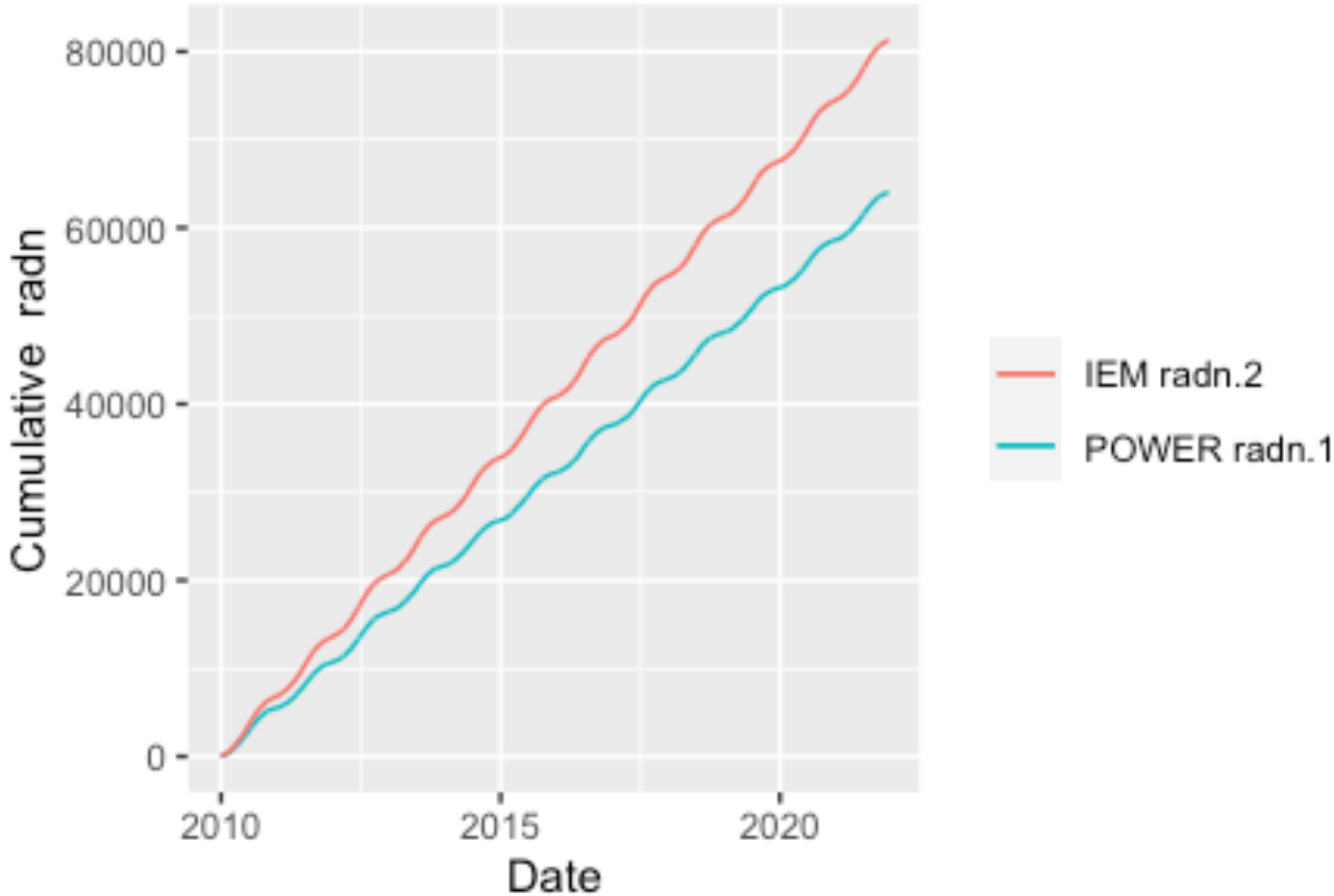
```r
library(apsimx)
## Get data close to the energy farm
pwr <- get_power_apsim_met(lonlat = c(-93.77, 42.02),
                           dates = c("2010-01-01","2021-12-10"))

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

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


IEM radn.2
30
20
10
0
10
20
30
POWER radn.1


80000
60000
40000
20000
0
Cumulative radn
2010
2015
2020
Date
IEM radn.2
POWER radn.1

APSIM Inputs: Weather and Soil

Obtaining weather data

One of the easiest ways of generating weather data (".met" files) is by using the 'nasapower' (https://github.com/ropensci/nasapower) or 'GSODR' (https://github.com/ropensci/GSODR) packages. For the US, some options are 'daymetr' (https://github.com/bluegreen-labs/daymetr) or through the Iowa Environmental Mesonet (https://mesonet.agron.iastate.edu/). In addition, this package provides the following utilities:

- **read_apsim_met**: read an APSIM met file into R with S3 class 'met'.
- **write_apsim_met**: write an APSIM met file from an object of class 'met'.
- **check_apsim_met**: check for reasonable values for met variables.
- **impute_apsim_met**: impute missing values using linear interpolation, splines or the mean.
- **napad_apsim_met**: Fill in date gaps with NAs in a 'met' object. Useful for imputation.
- **compare_apsim_met**: utility for quick comparison of 'met' objects. There is also a plot method.
- **get_power_apsim_met**: (requires the 'nasapower' package). Fetch data and create a 'met' R object and (optionally) write a met file to disk.
- **get_daymet2_apsim_met**: (requires the 'daymetr' package). Fetch data and create a 'met' R object and (optionally) write a met file to disk.
- **get_gsod_apsim_met**: (requires the 'GSODR' package). Fetch data and create a 'met' R object and (optionally) write a met file to disk.
- **get_iem_apsim_met**: Fetch data from Iowa Environmental Mesonet
- **get_iemre_apsim_met**: Fetch data from Iowa Environmental Mesonet Reanalysis

https://femiguez.github.io/apsimx-docs/apsimx.html#apsim-inputs-weather-and-soil

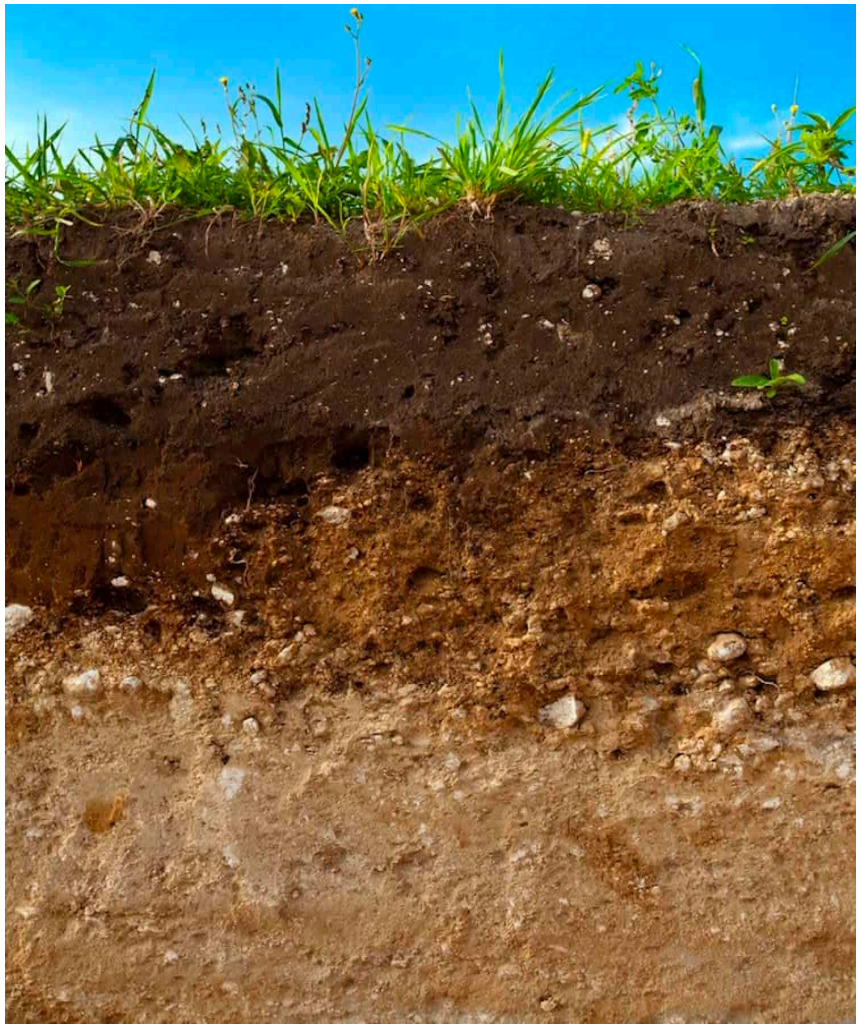
Outline

- Sources of weather data
- R activities
- Sources of soil data
- R activities
- Additional Resources

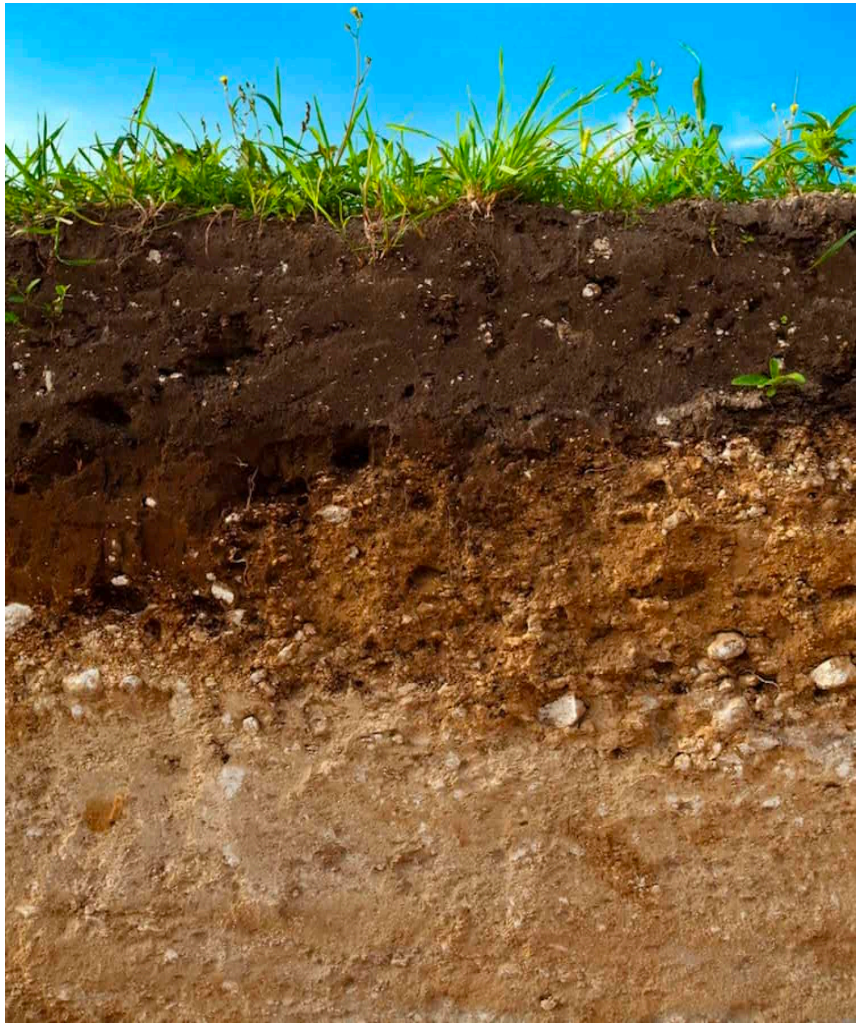
Outline

- Sources of weather data
- R activities
- Sources of soil data
- R activities
- Additional Resources

Let’s talk about soils…

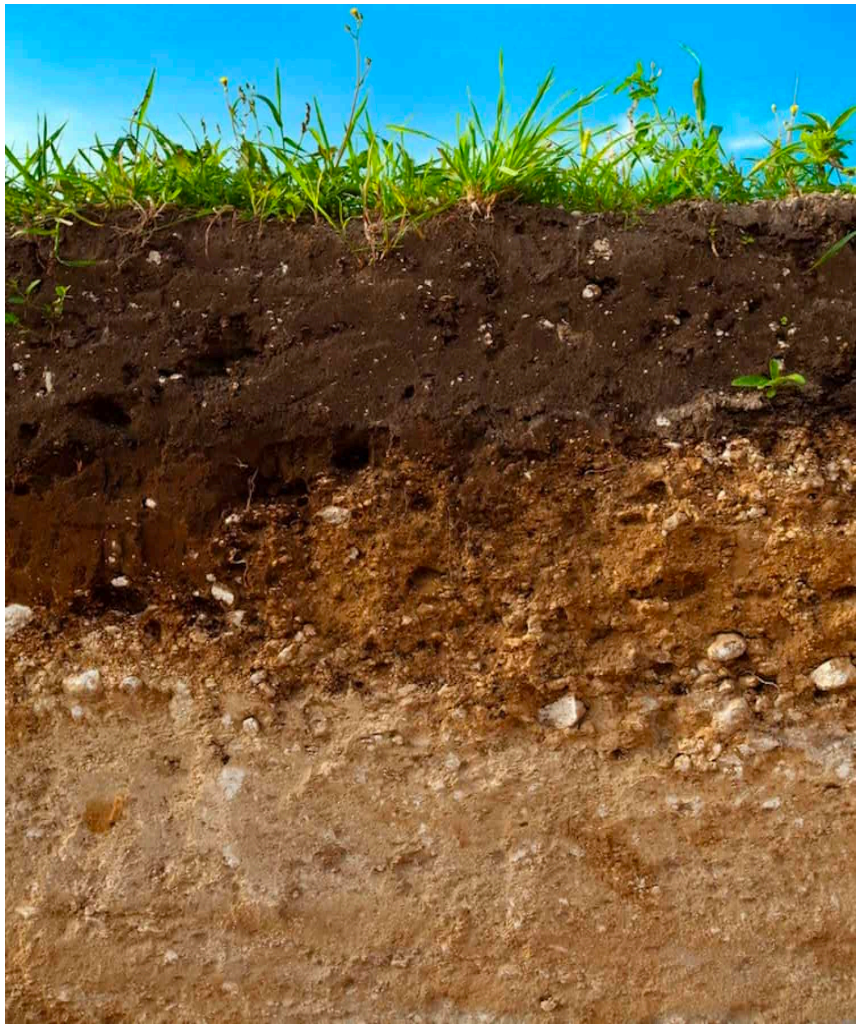


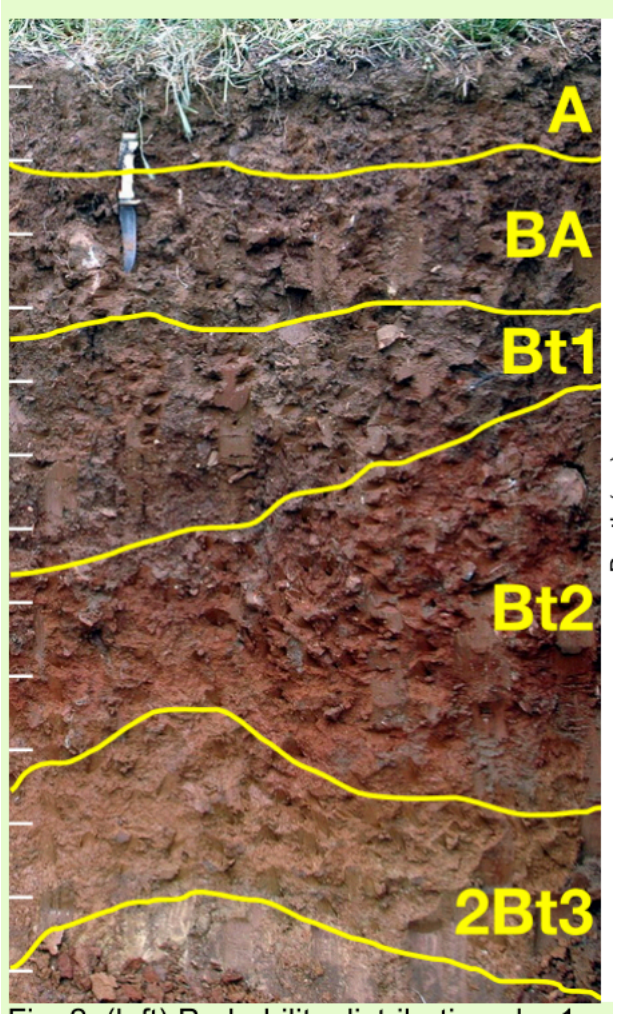
Let’s talk about soils…



- Not the same as weather data
- Fewer sources of data
- More spatially variable
- Need to consider depth
- Soil sampling is costly

Soils and Soil Database





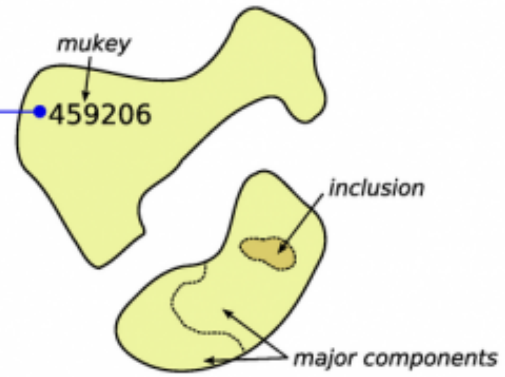
Soil Databases

- SSURGO (USDA-NRCS) / STATSGO
- gridded SSURGO
- SoilGrid - ISRIC
- Harmonized World Soil Database
- FAO soils
- Polaris

Soil Databases

- SSURGO (USDA-NRCS) / STATSGO
- gridded SSURGO
- SoilGrid - ISRIC
- Harmonized World Soil Database
- FAO soils
- Polaris

SSURGO Table Diagram



Map Unit Polygons

(components are unmapped)

mapunit: *mukey*

459206 ...

mukey **:component:** *cokey*

459206 ... 459206:623924 ***inclusion***

459206 ... 459206:623925 ***major component***

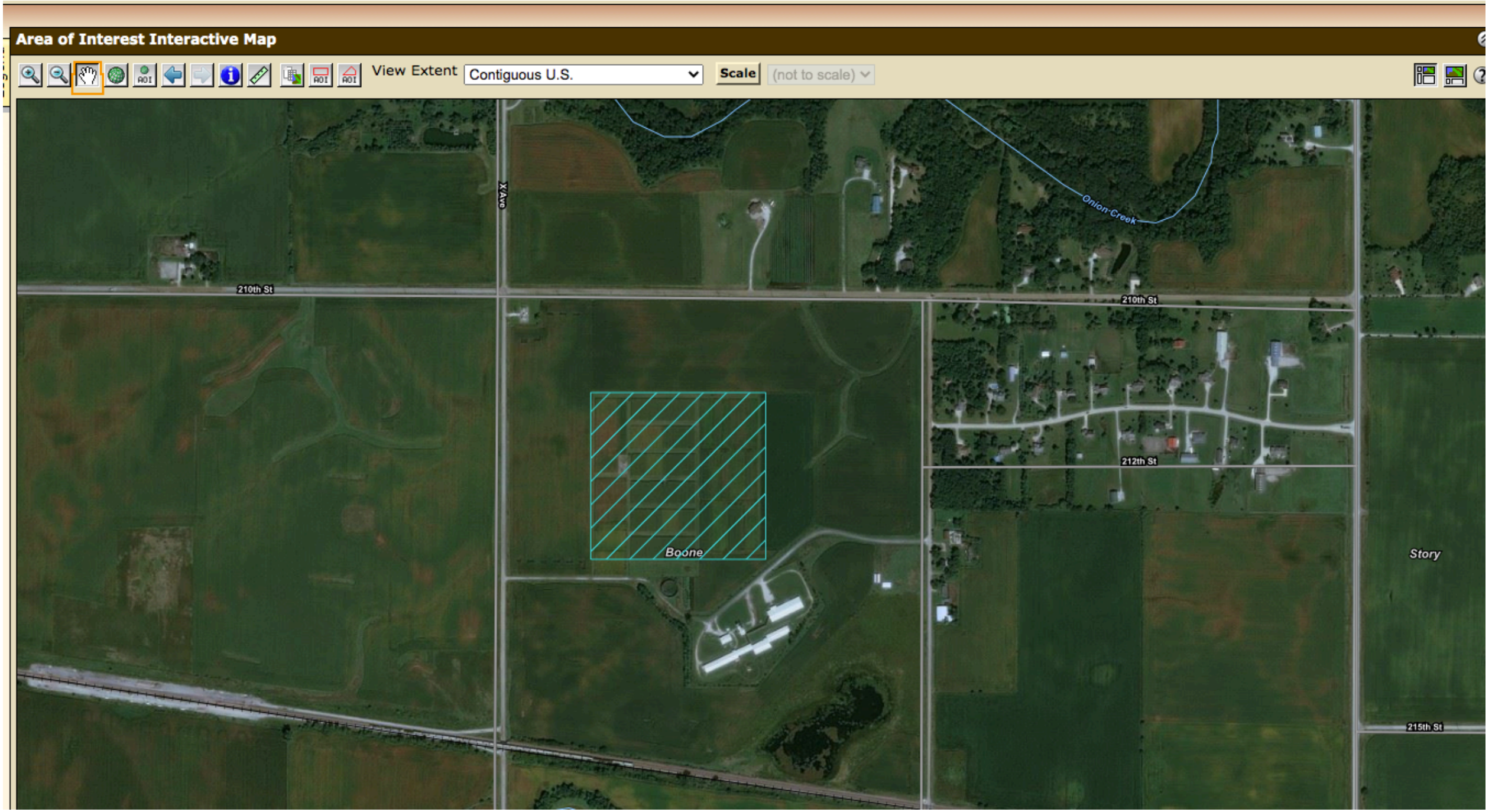
cokey **:chorizon:** *hznum*

459206:623925 ... *1*

459206:623925 ... 2

459206:623925 ... 3

https://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm


Area of Interest Interactive Map
View Extent
Contiguous U.S.
Scale
(not to scale)
X Ave
Onion Creek
210th St
210th St
212th St
Boone
Story
215th St

Tables — Organic Matter — Summary By Map Unit

Summary by Map Unit — Boone County, Iowa (IA015)

Summary by Map Unit — Boone County, Iowa (IA015)

Map unit symbol	Map unit name	Rating (percent)	Acres in AOI	Percent of AOI
6	Okoboji silty clay loam, 0 to 1 percent slopes	8.00	0.0	0.1%
L55	Nicollet loam, 1 to 3 percent slopes	5.33	4.9	19.3%
L95	Harps clay loam, Bemis moraine, 0 to 2 percent slopes	6.33	0.7	2.8%
L138B	Clarion loam, Bemis moraine, 2 to 6 percent slopes	3.27	3.3	13.0%
L138C	Clarion loam, Bemis moraine, 6 to 10 percent slopes	3.17	1.0	3.9%
L138C2	Clarion loam, Bemis moraine, 6 to 10 percent slopes, moderately eroded	1.63	0.1	0.2%
L507	Canisteo clay loam, Bemis moraine, 0 to 2 percent slopes	6.33	15.5	60.8%
Totals for Area of Interest			**25.5**	**100.0%**

https://en.wikipedia.org/wiki/USDA_soil_taxonomy
https://soilseries.sc.egov.usda.gov/osdname.aspx

Area of Interest (AOI) | Soil Map | Soil Data Explorer | **Download Soils Data** | Shopping Cart (Free)

Download Soils Data for...

Your AOI (SSURGO)

Create Download Link

General Information

Link	Description of Soil Survey Geographic (SSURGO) Database
Download Contents	Tabular data, spatial data (if available), thematic map data, template database, and FGDC metadata
Spatial Data Format	ESRI Shapefile, Geographic WGS84

Soils Data Download Package for your AOI (SSURGO)

AOI Location
Boone County, Iowa

Soil Survey Areas

Boone County, Iowa (IA015)

Area in AOI
25.5 acres

Data Availability
Tabular and Spatial, complete

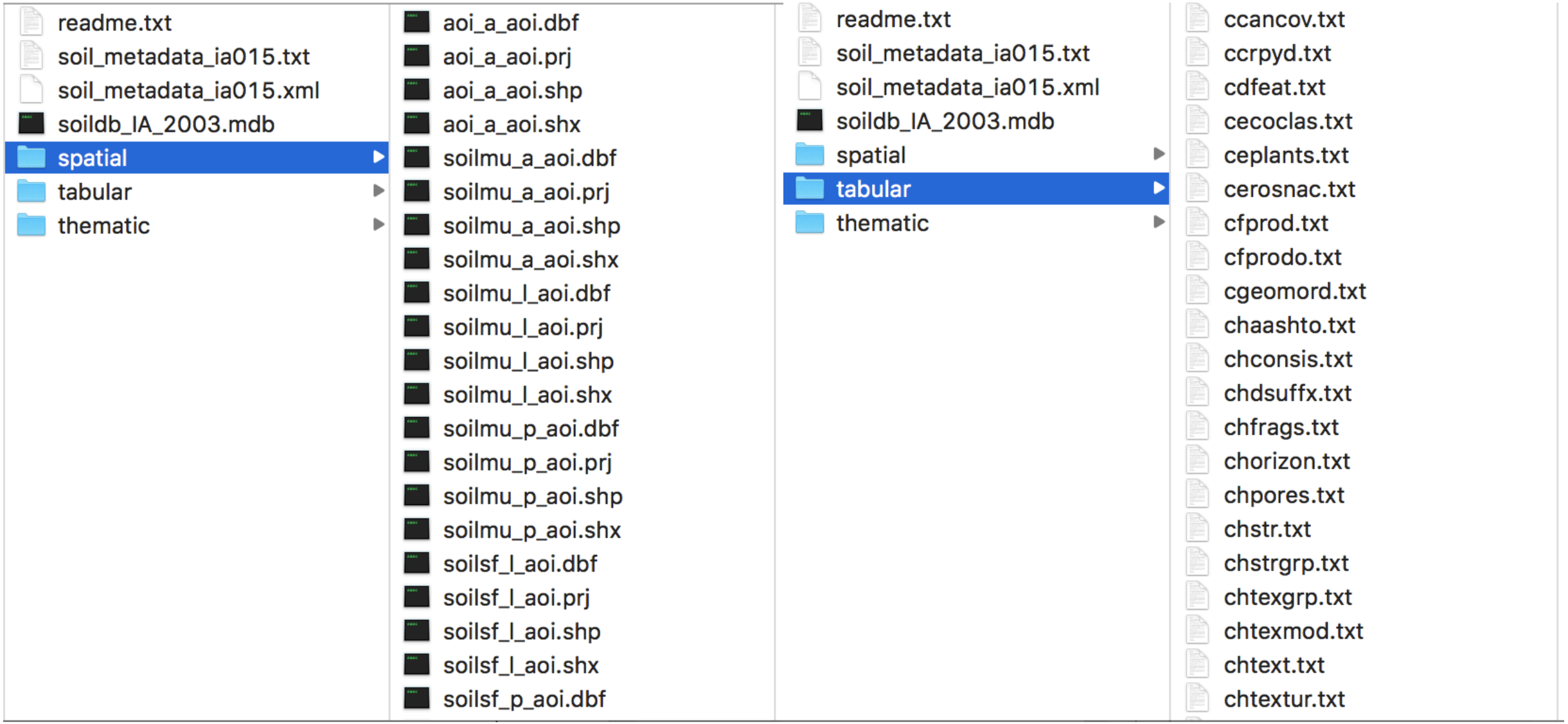
Version
Survey Area: Version 28, Jun 10, 2020
Tabular: Version 27, Jun 10, 2020
Spatial: Version 11, Sep 13, 2019

Template Database
State: IA
Microsoft Access Version: Access 2003
Template Database Version: 36
Template Database Name: soildb_IA_2003

Download Size
3.2 MB

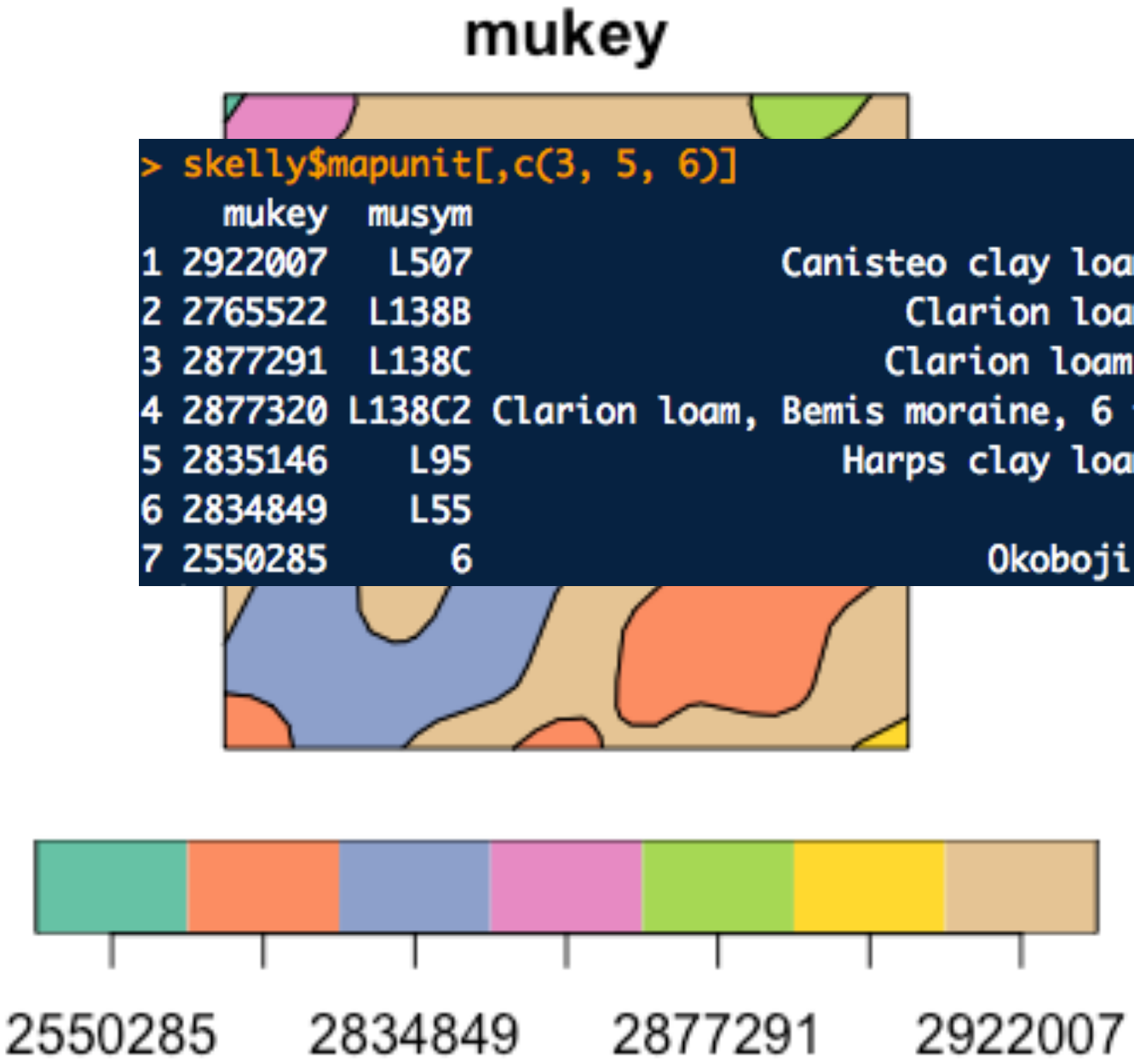
Download Link
wss_aoi_2021-09-16_09-12-37.zip

Create Download Link



29 “spatial” files

71 “txt” files




```
> skelly$mapunit[,c(3, 5, 6)]
     mukey  musym                                                                muname
1  2922007   L507                    Canisteo clay loam, Bemis moraine, 0 to 2 percent slopes
2  2765522  L138B                         Clarion loam, Bemis moraine, 2 to 6 percent slopes
3  2877291  L138C                        Clarion loam, Bemis moraine, 6 to 10 percent slopes
4  2877320 L138C2 Clarion loam, Bemis moraine, 6 to 10 percent slopes, moderately eroded
5  2835146    L95                       Harps clay loam, Bemis moraine, 0 to 2 percent slopes
6  2834849    L55                                    Nicollet loam, 1 to 3 percent slopes
7  2550285      6                      Okoboji silty clay loam, 0 to 1 percent slopes
```

Outline

- Sources of weather data
- R activities
- Sources of soil data
- R activities
  - Retrieving
  - Visualizing
  - Comparing
- Additional Resources

Soil Data with R

- soilDB package, you might need other packages

```
## 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)
class(ams.tbls)
ams.tbls$mapunit.shp
```

Soil Data with R

```
> ## 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

```
## Retrieving an area
ams.tbls2 <- get_ssurgo_tables(lonlat = c(-93.77, 42.02), shift = 300)
ams.tbls2$mapunit.shp$mukey <- as.factor(ams.tbls2$mapunit.shp$mukey)
plot(ams.tbls2$mapunit.shp[, "mukey"], key.pos = 1)
```

Retrieve soil profile data and return a table with data

Description

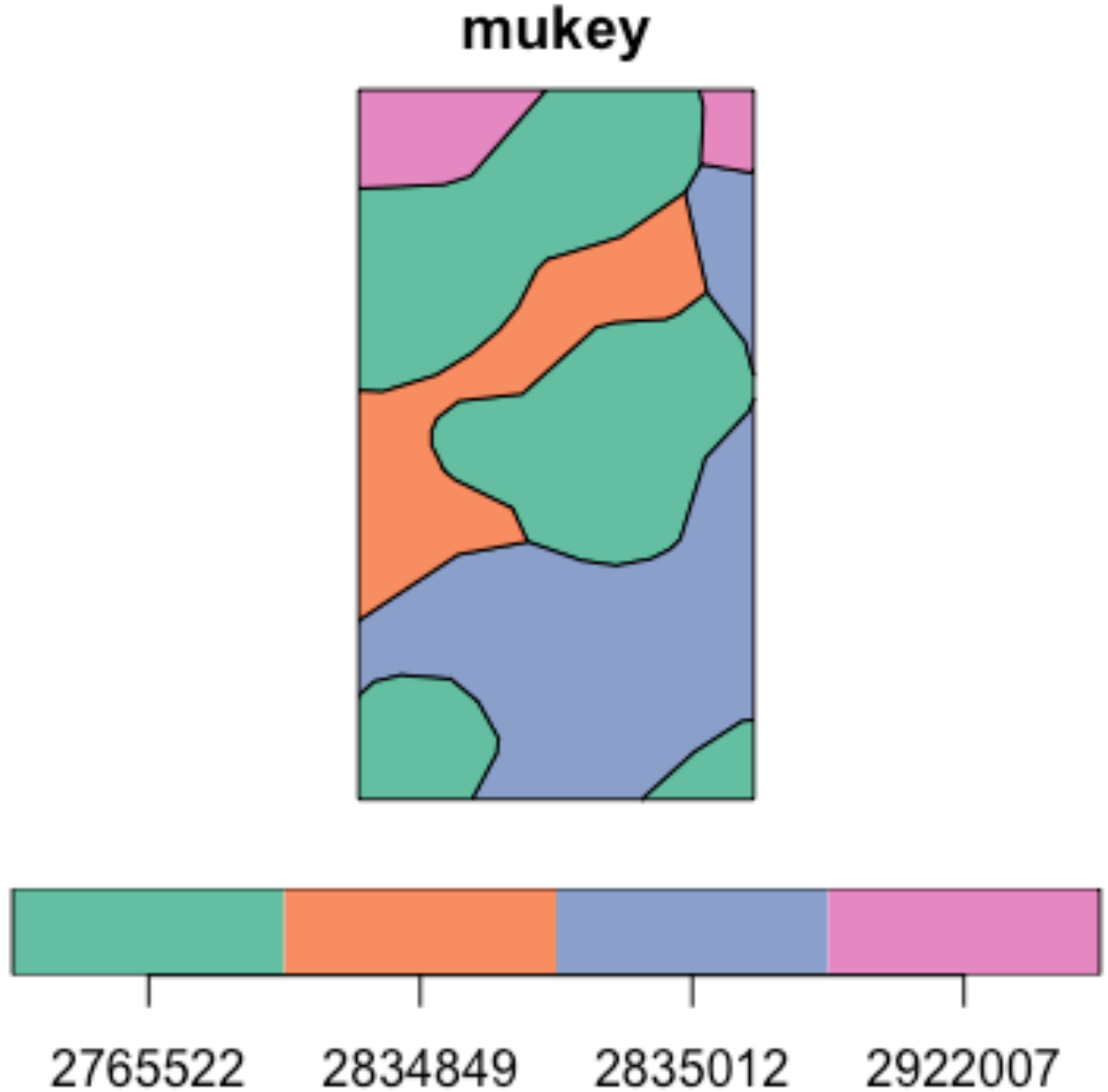
This function does partially what get_ssurgo_soil_profile does, but it returns a list with tables for mapunit, component, chorizon and mapunit.shp (object of class sf)

Usage

```
get_ssurgo_tables(lonlat, shift = -1, aoi, verbose = FALSE)
```

Arguments

`lonlat`	Longitude and latitude vector (e.g. c(-93, 42))
`shift`	simple mechanism for creating an area of interest by displacing the point indicated in lonlat by some amount of distance (e.g. 300 - in meters)
`aoi`	area of interest, if supplied the lonlat and shift arguments will be ignored. Should be of class 'sp::SpatialPolygons'.
`verbose`	whether to pring messages and warnings to the console default FALSE


mukey
2765522
2834849
2835012
2922007

Getting soil profiles

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

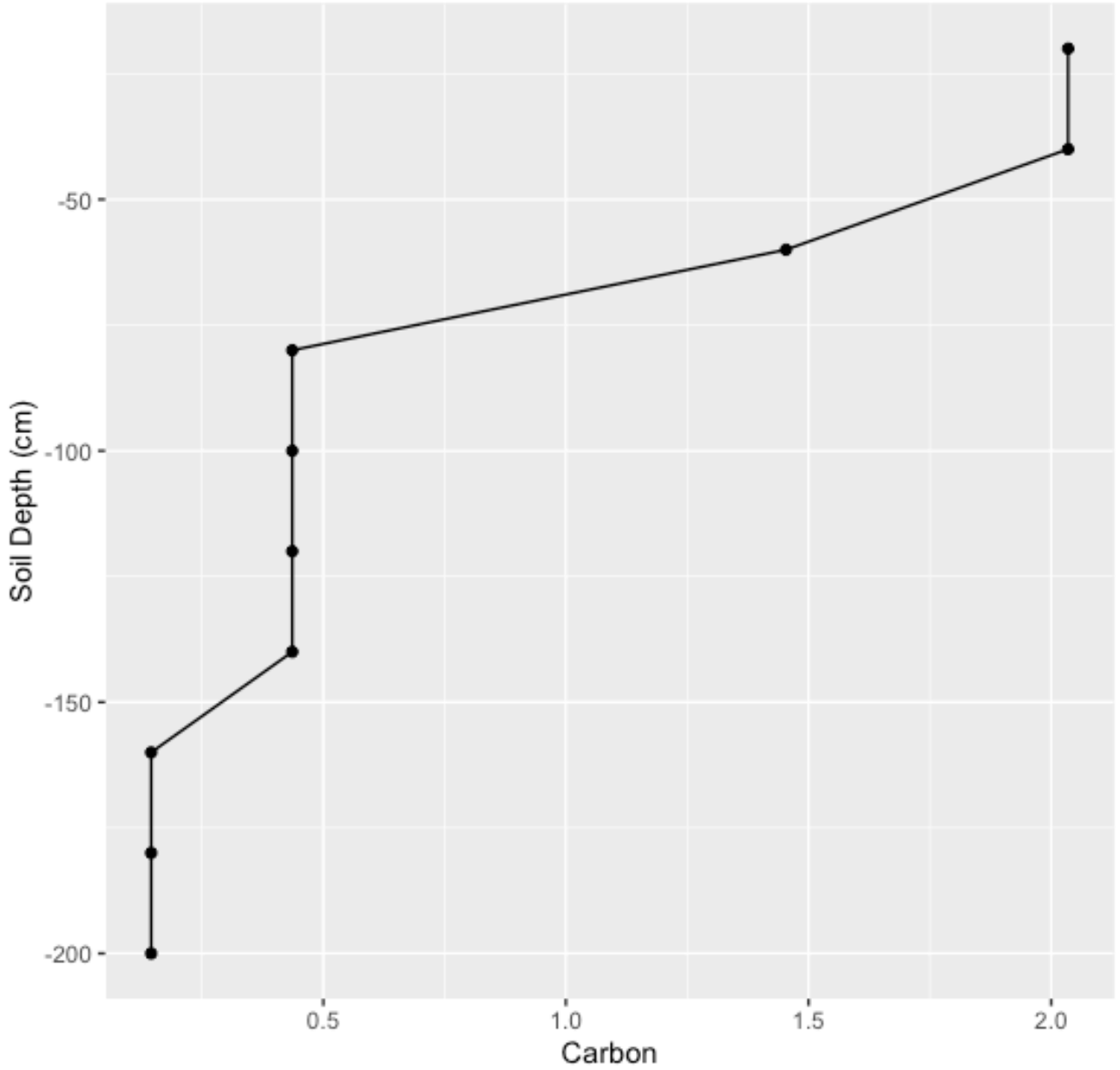
```
> sps[[1]]$metadata$SoilType
[1] "Clarion:2765522"
> sps[[2]]$metadata$SoilType
[1] "Storden:2765522"
>
```

Outline

- Sources of weather data
- R activities
- Sources of soil data
- R activities
  - Retrieving
  - Visualizing
  - Comparing
- Additional Resources

Soil Carbon

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



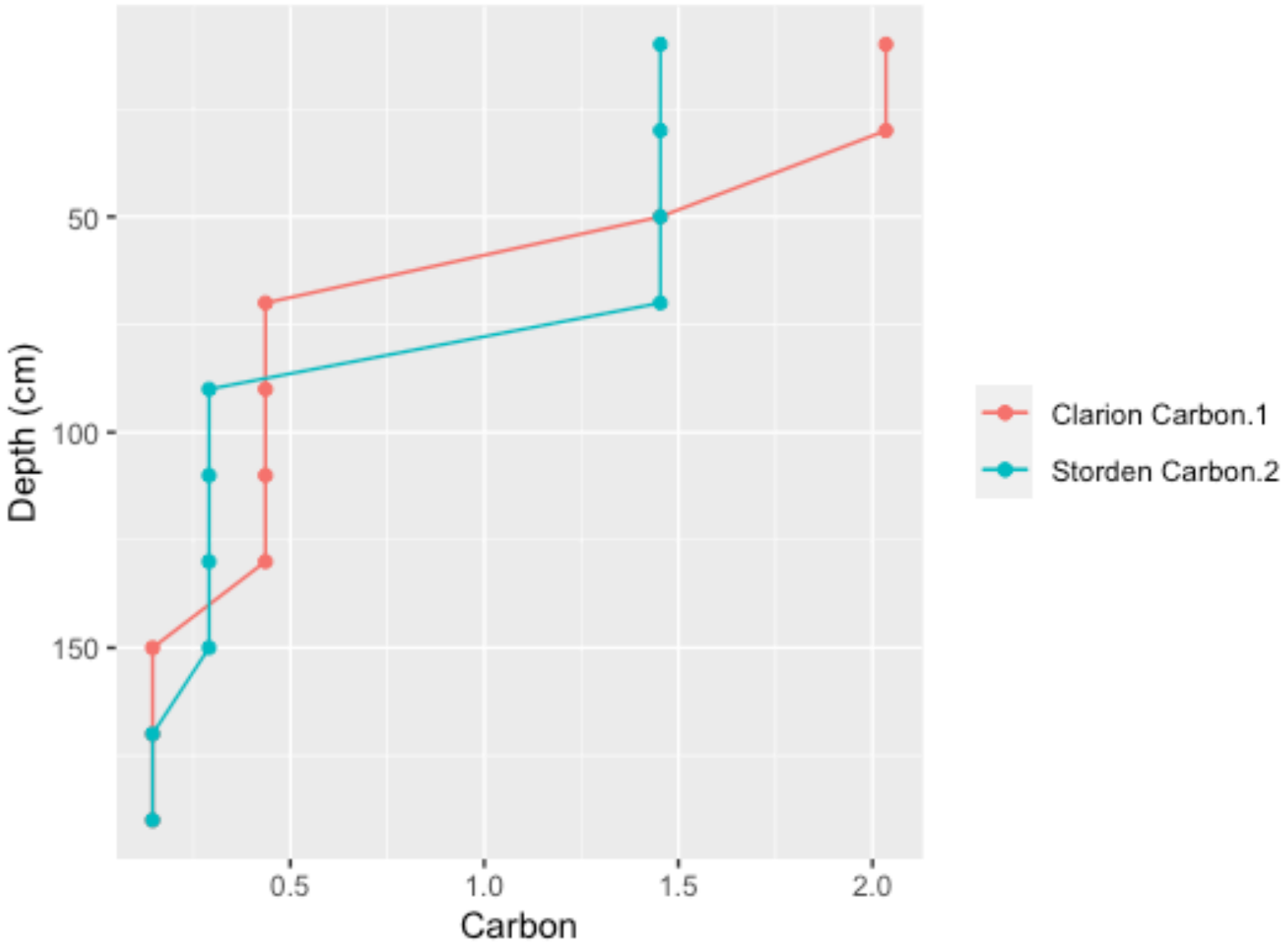
Outline

- Sources of weather data
- R activities
- Sources of soil data
- R activities
  - Retrieving
  - Visualizing
  - Comparing
- Additional Resources

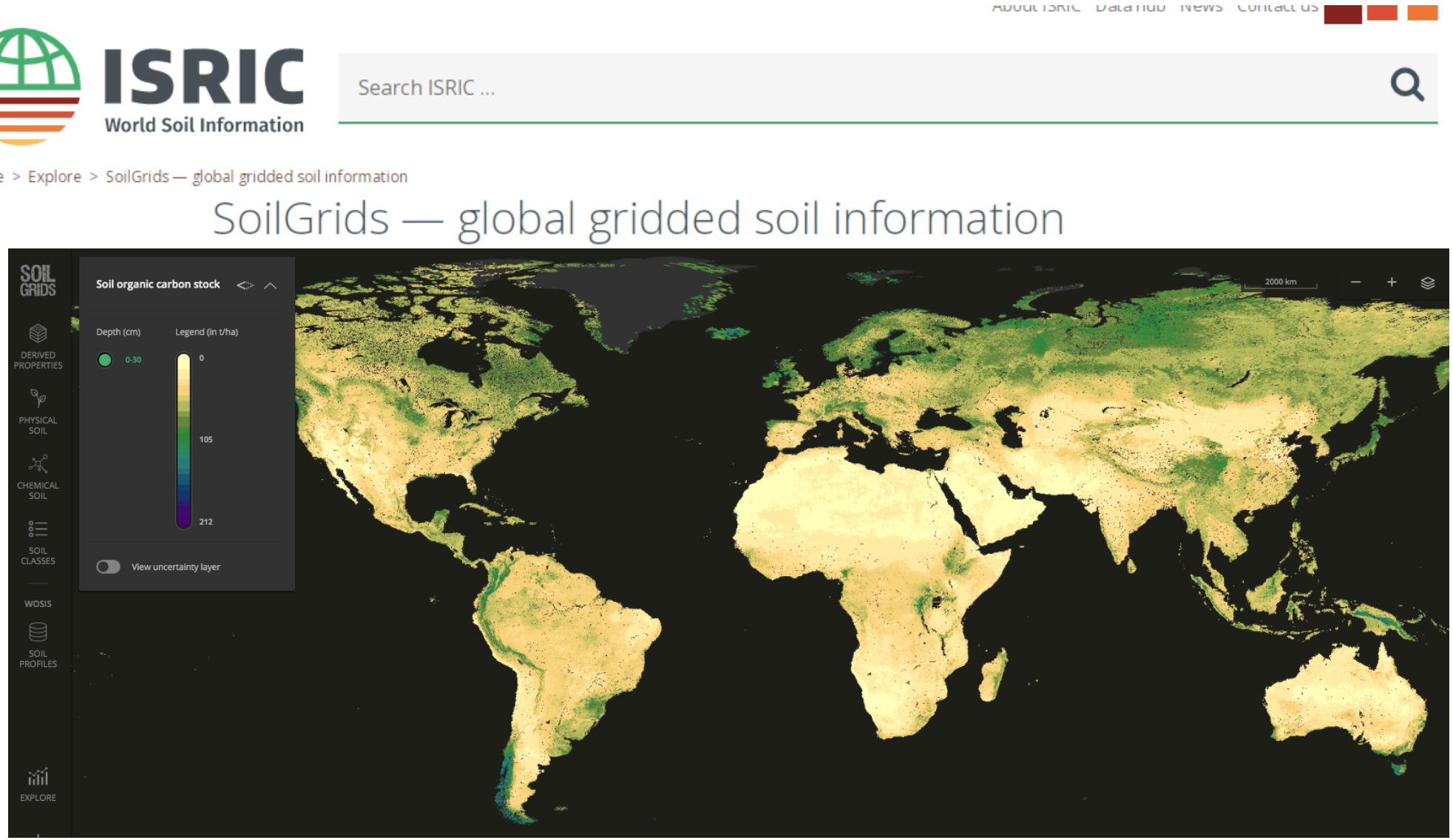
Soil Carbon

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

Soil Carbon

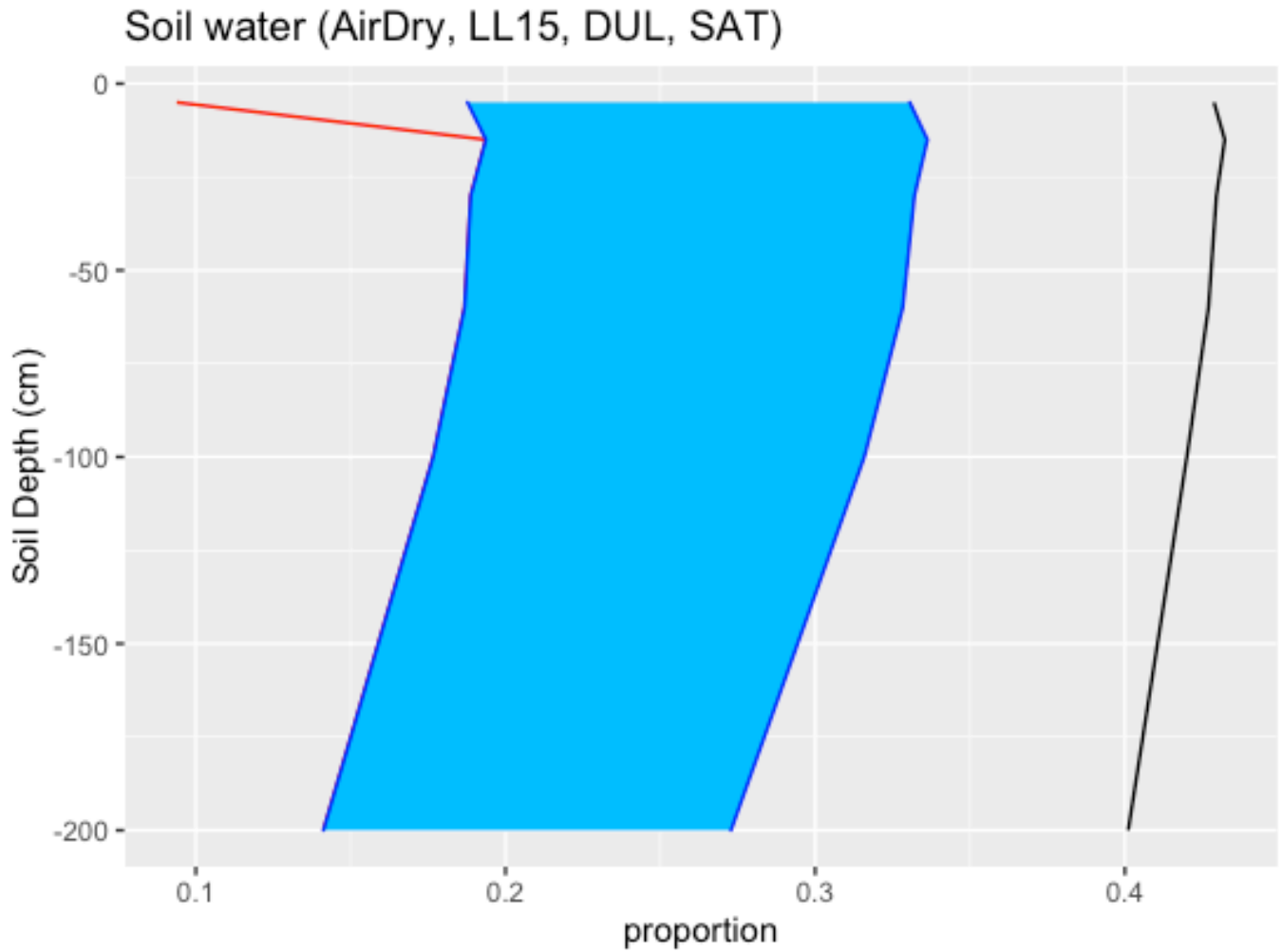

Depth (cm)
50
100
150
0.5
1.0
1.5
2.0
Carbon
Clarion Carbon.1
Storden Carbon.2

Global Soil Databases



https://soilgrids.org

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

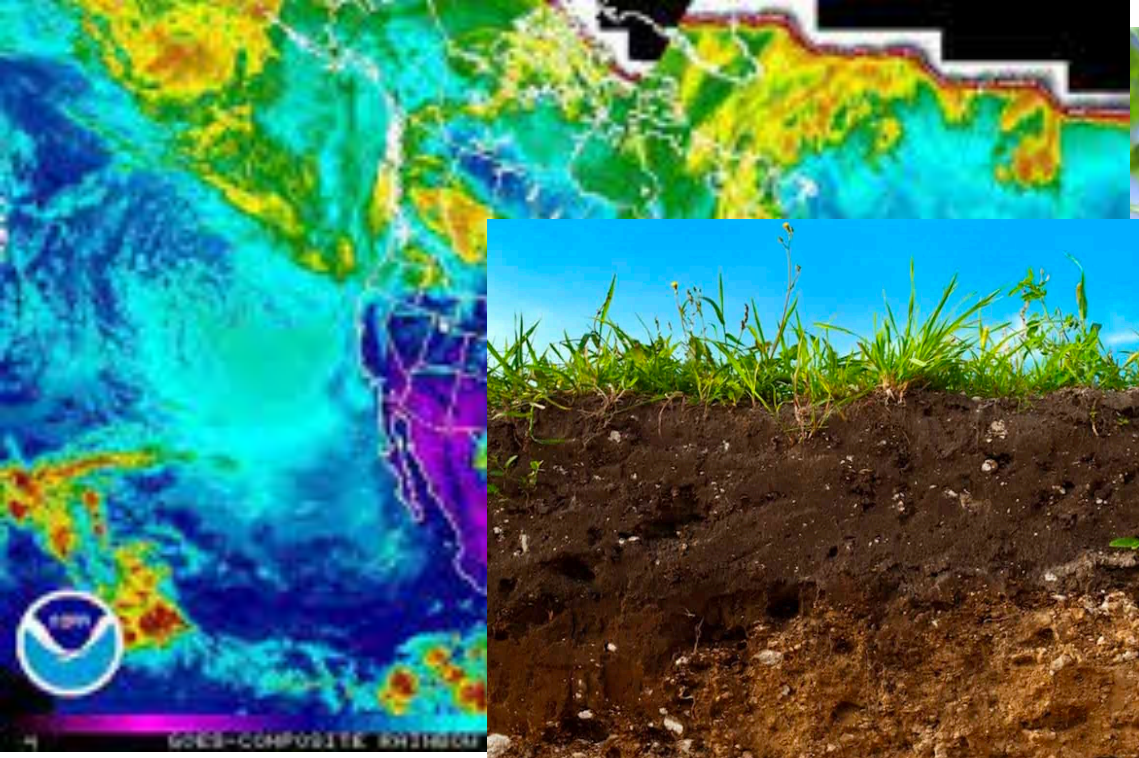


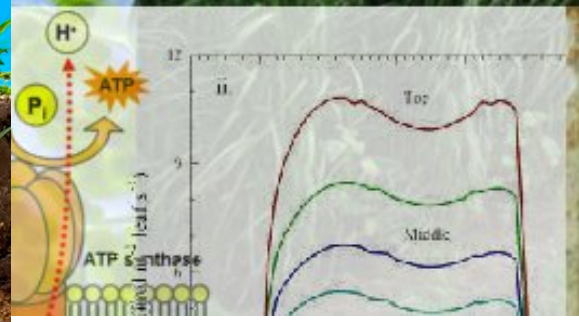
Outline

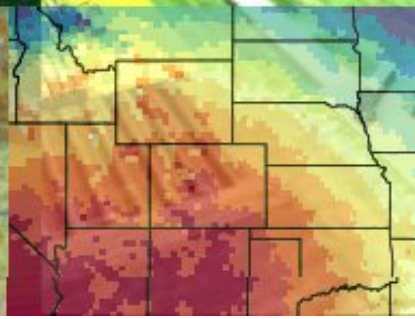
- Sources of weather data
- R activities
- Sources of soil data
- R activities
- Additional Resources

Additional resources

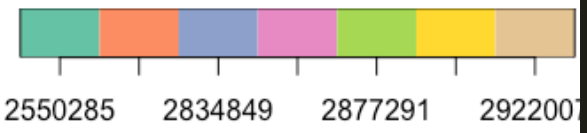
- APSIM Initiative Training Videos
  - Calculating Thermal Time for Use in APSIM
  - Understanding APSIM Next Generation weather (Matthew Harrison)
  - Soil, Weather, Experiments, Calibration
- Publications
  - A protocol to build soil descriptions for APSIM simulations (Cichota et al. 2021)
- Other
  - https://www.apsim.info/
  - https://apsimnextgeneration.netlify.app/
  - Software (Python): bestiapop (https://github.com/JJguri/bestiapop)



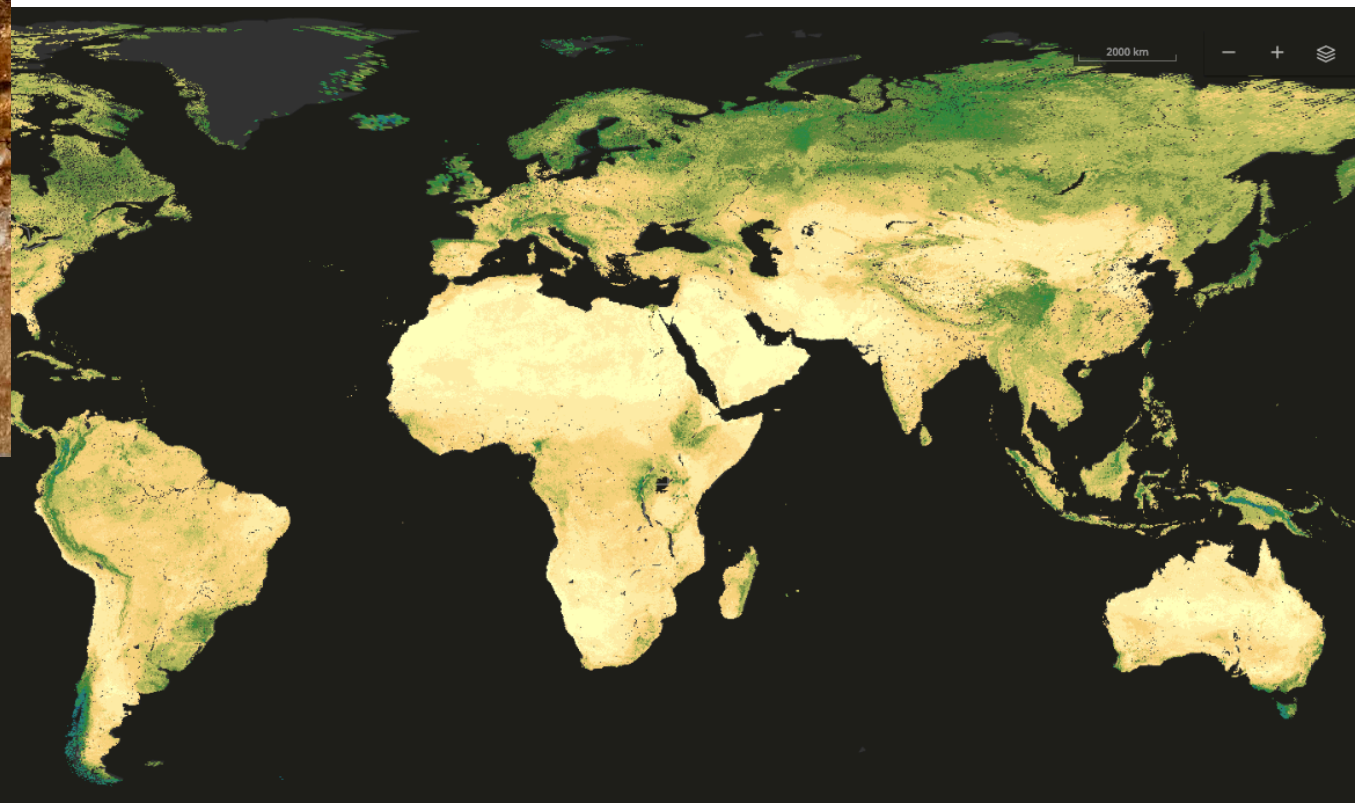

H+
ATP
Pi
ATP synthase
Top
Middle




mukey


2550285
2834849
2877291


WOSIS
SOIL PROFILES
EXPLORE


2000 km