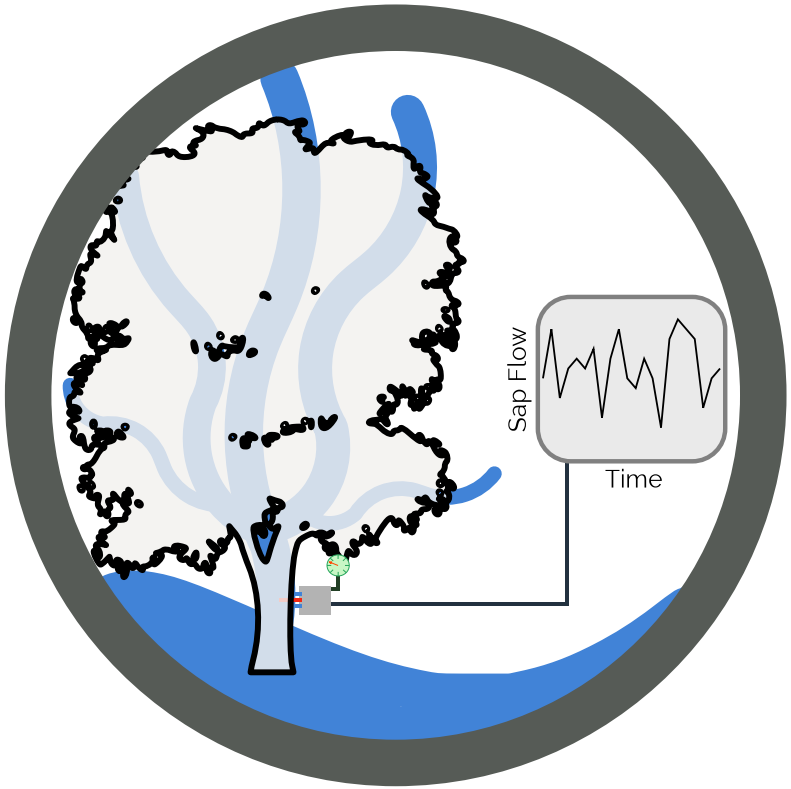
Compiling a global database of sap flow measurements: the SAPFLUXNET data workflow

XIV MEDECOS & XIII AEET meeting
Ecoinformatics: data science brings new avenues for ecology Symposium

Víctor Granda, Rafael Poyatos, Roberto Molowny-Horas, Maurizio Mencuccini, Kathy Steppe & Jordi Martínez-Vilalta

CREAF
Centre of Ecological Research and Forestry Applications

Introduction



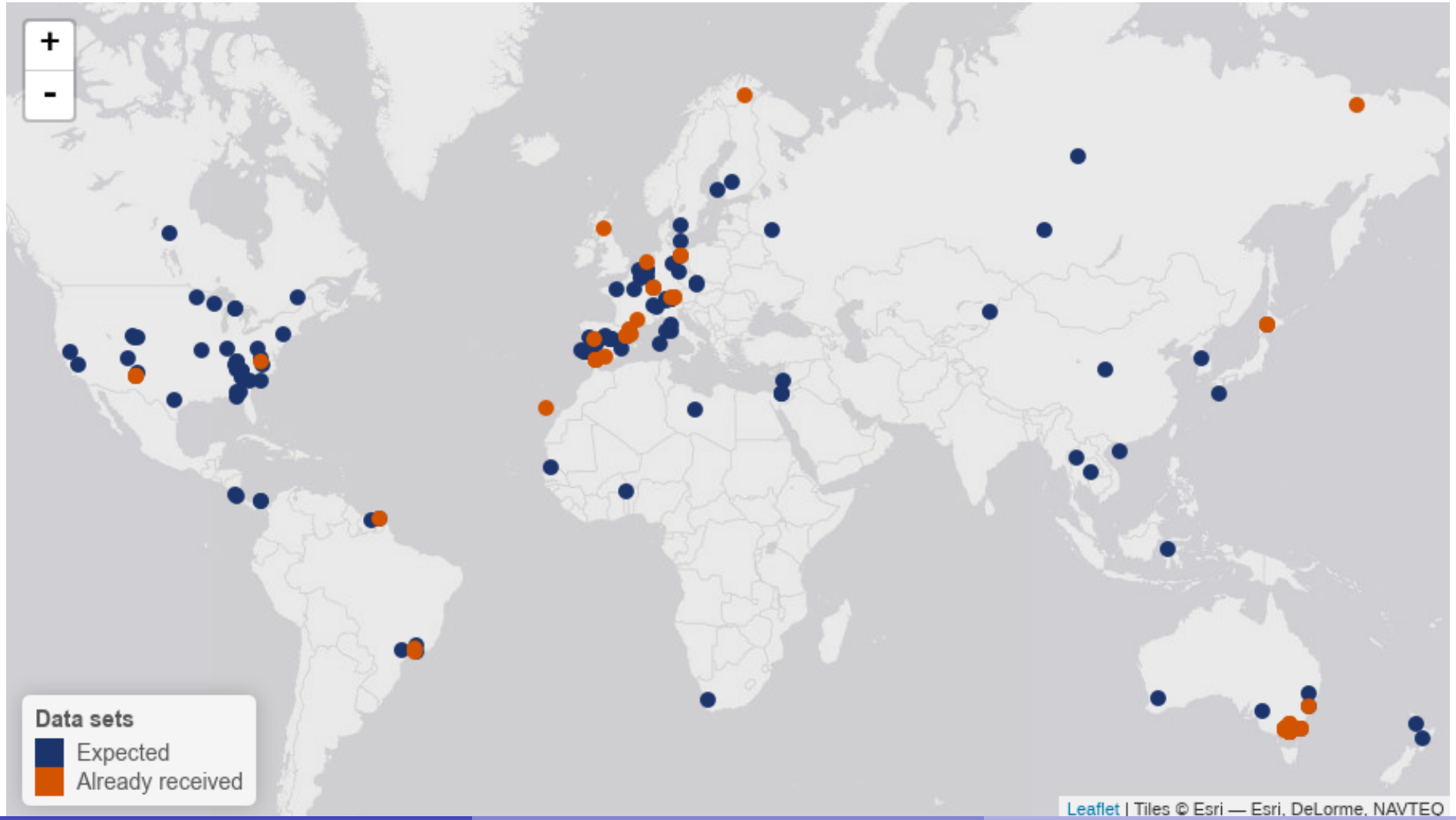
Proxy of the movement of water between the soil-plant-atmosphere continuum.

Different **thermodynamic methods** to determine sap flow using heat as a tracer sap movement

Allows **upscaling** from stem to plant and landscape level.

The time is ripe for a global database

The **SAPFLUXNET** initiative is building the first global database of plant-level sap flow measurements to analyse the environmental and physiologycal factors driving tree- and stand-level transpiration



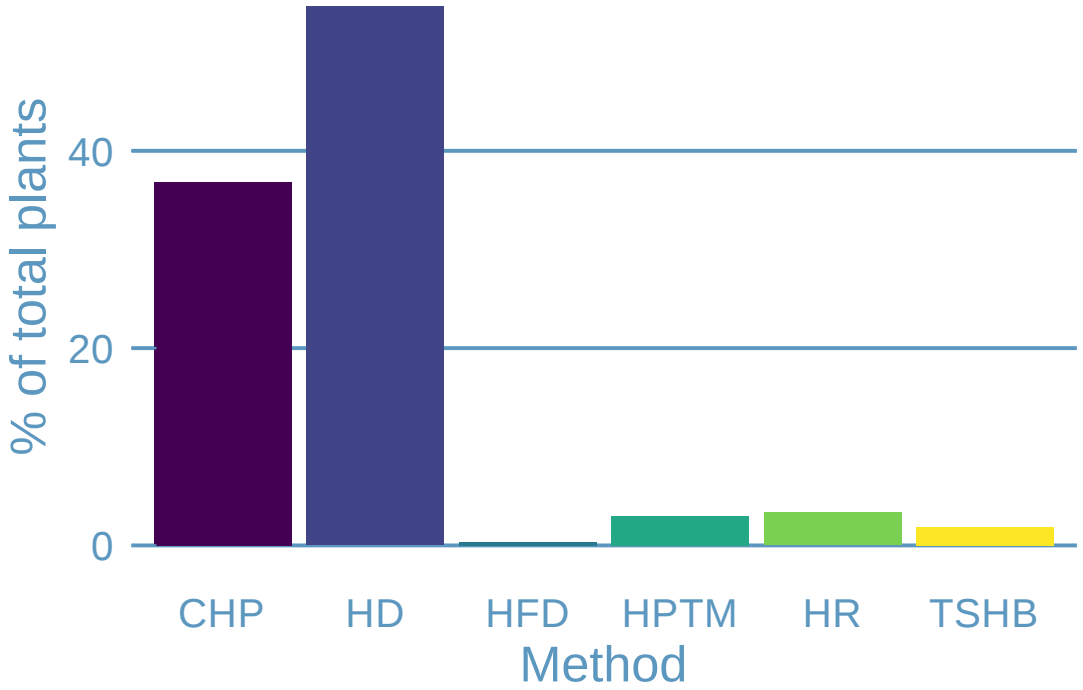
Target datasets

- Stem or whole-plant level
- Field conditions
- Sub-daily intervals
- Environmental data available (RH, Ta, PAR...)
- Abundant metadata (site, stand, plant, species and environmental)

Data Characteristics

High data complexity:

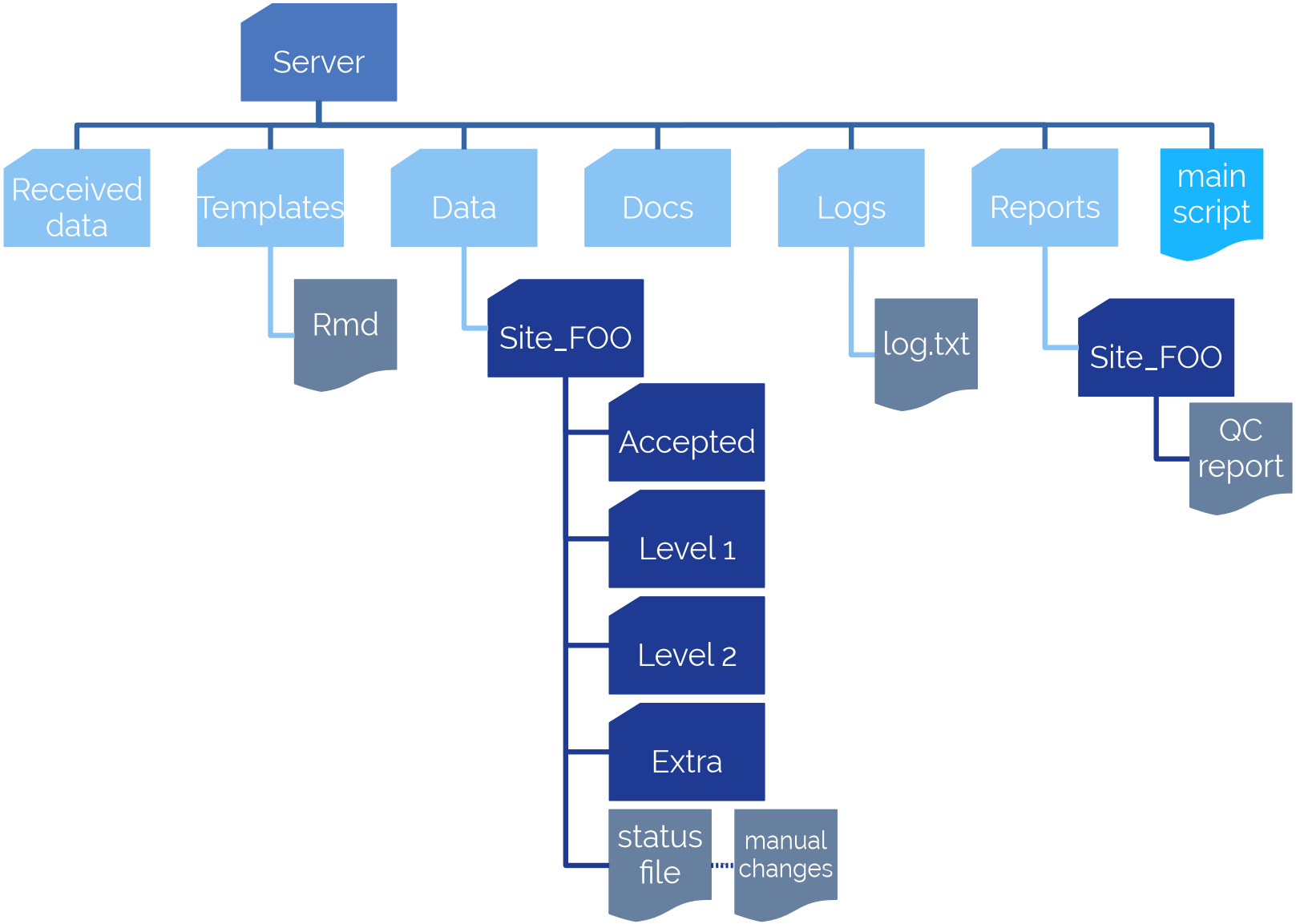
Metadata	Items
Site	20
Stand	16
Species	4
Plant	24
Environmental	16
Total	**80**



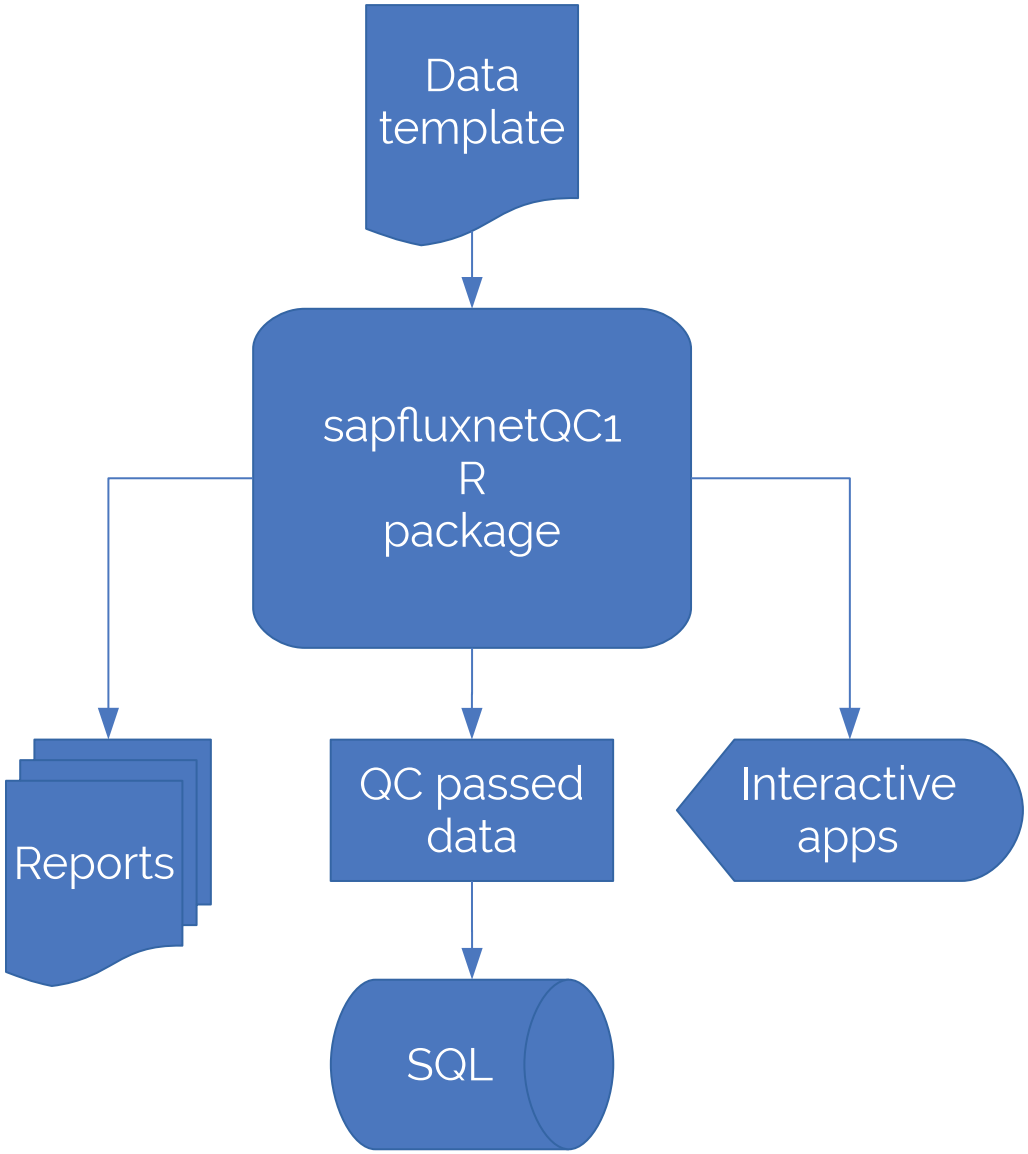
How to assay quality and store data?

We need **semi-automatic**, **reproducible** and **robust** checks to ensure the quality of the submmited datasets. Also, we need to store the data in a way that allows all essential information to be **available** in order to use the data in checks and analyses

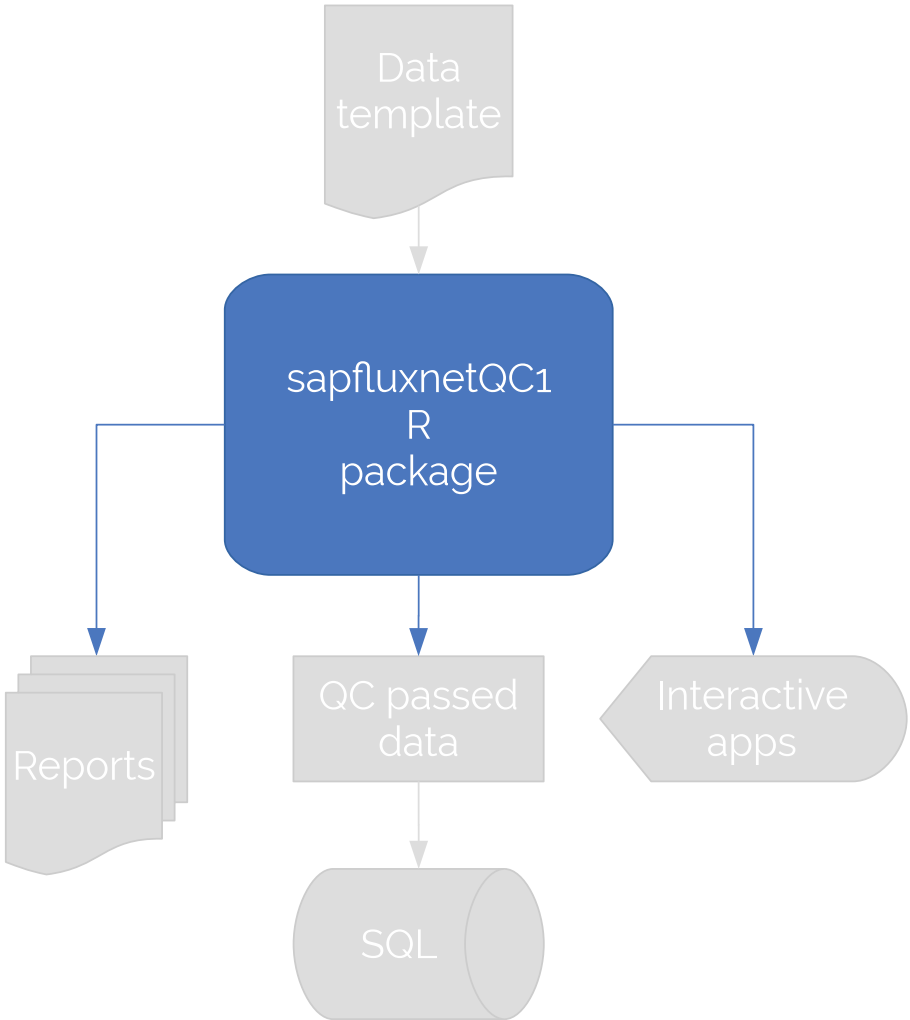
SAPFLUXNET Server Structure



SAPFLUXNET Work Flow

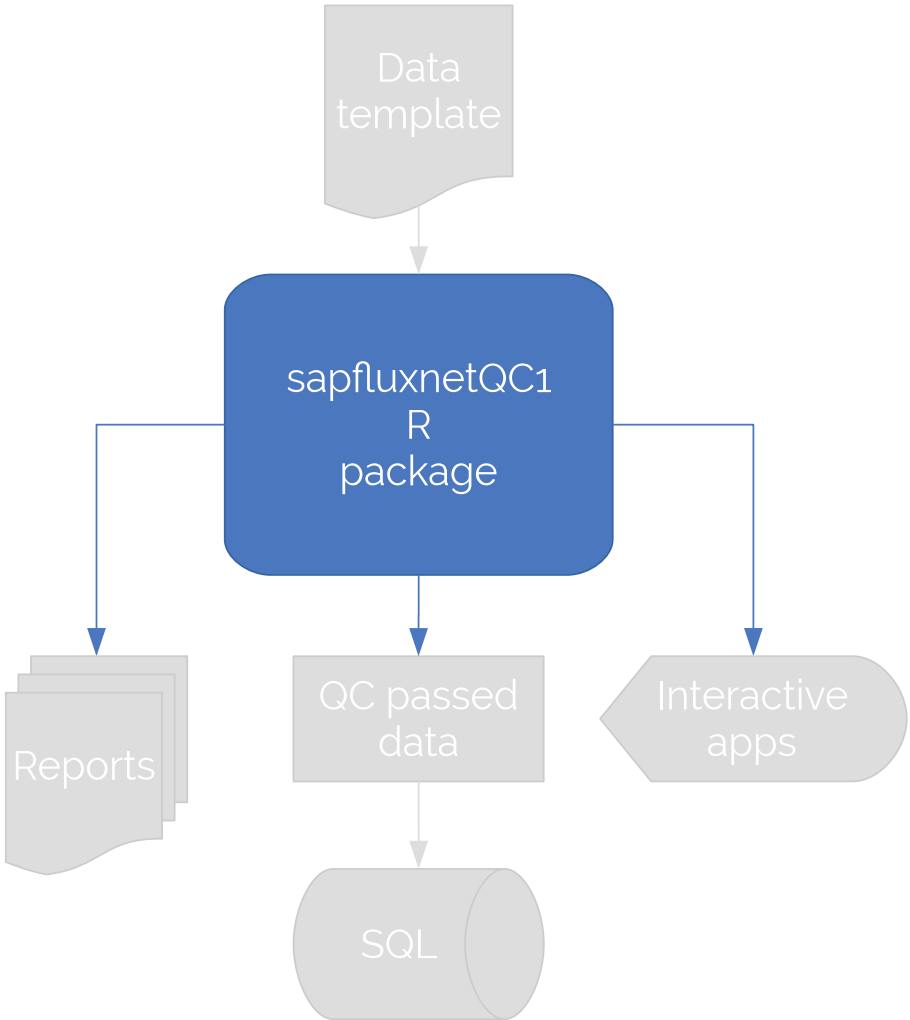


SAPFLUXNET Work Flow



- Automatic Quality Control checks (QC)
- Automatic report generation
- Storing data in special object (SfnData S4 class)
- Interactive functions allowing fine control of QC

SAPFLUXNET Work Flow



Benefits of R as development environment

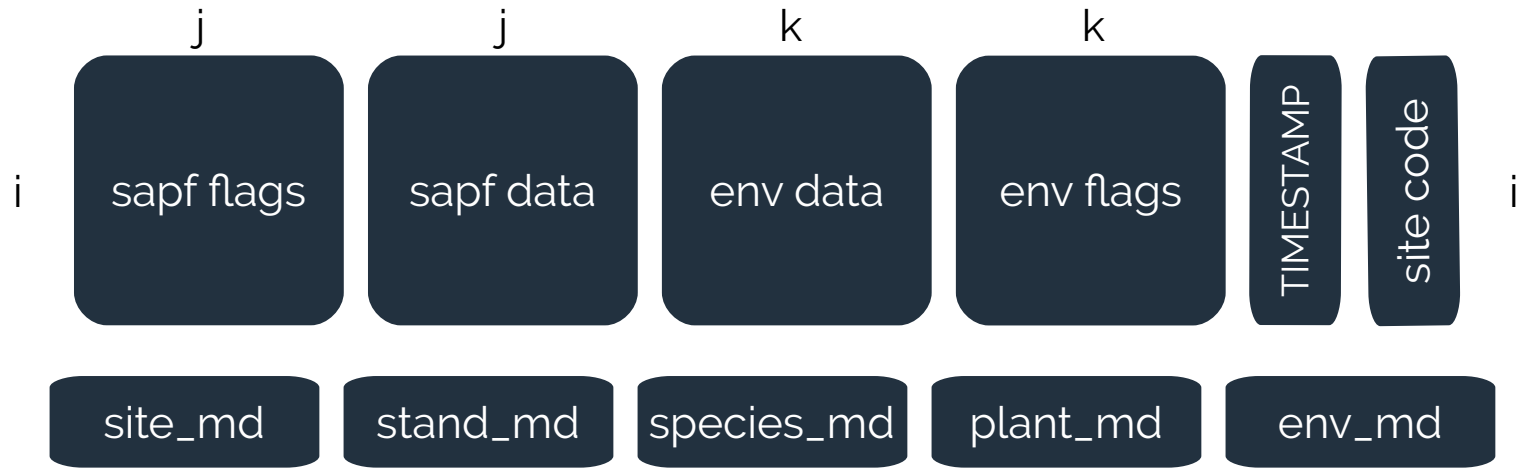
- Open
- Reproducible
- Easy maintenance and update
- Easy integration with web and SQL technologies

SAPFLUXNET SfnData S4 Class

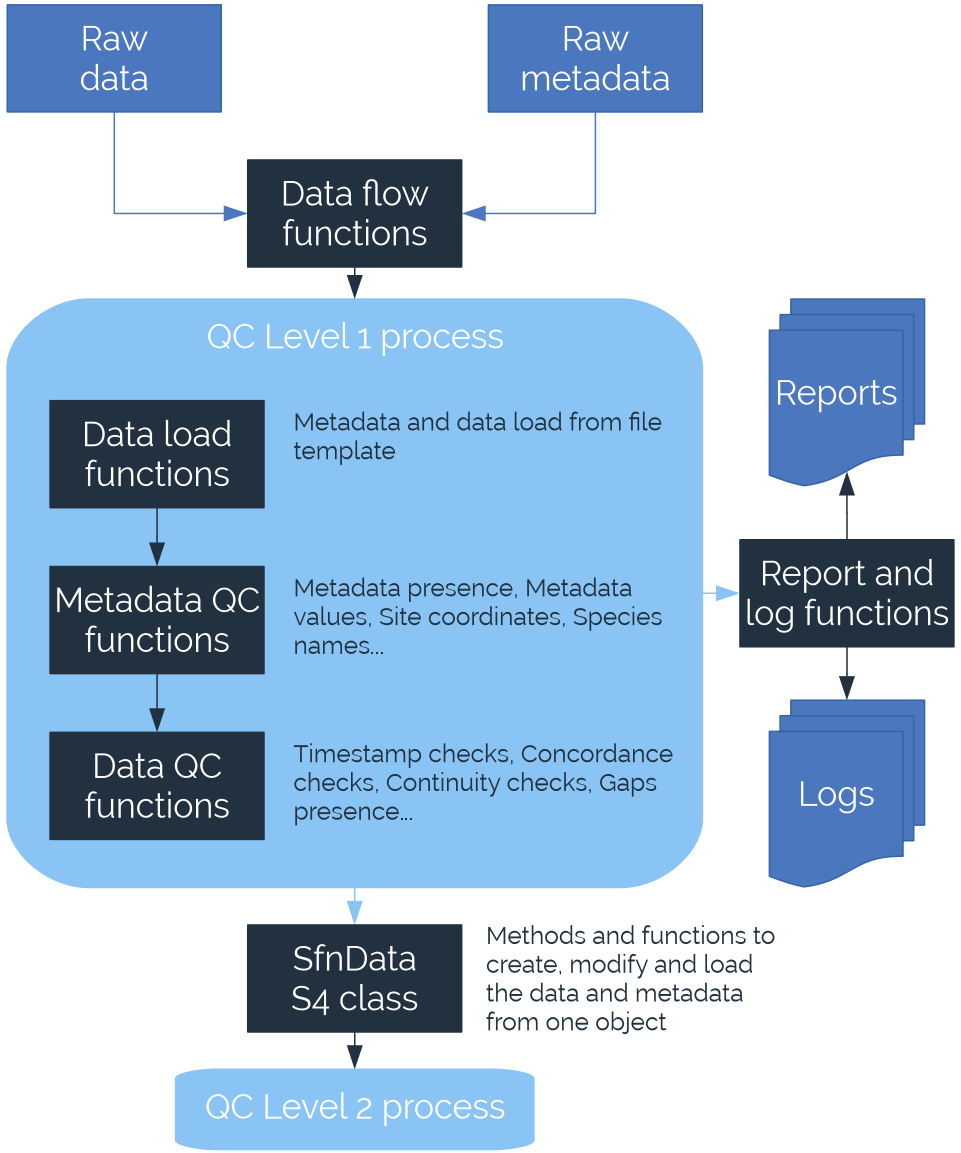
S4 classes:

- Easy way of store complete site data
- Validity checks based in fair assumptions
- Methods: `[]`, `get`, `<-`
- Scalable: allows for combining sites for more complex analyses (**whishlist**)

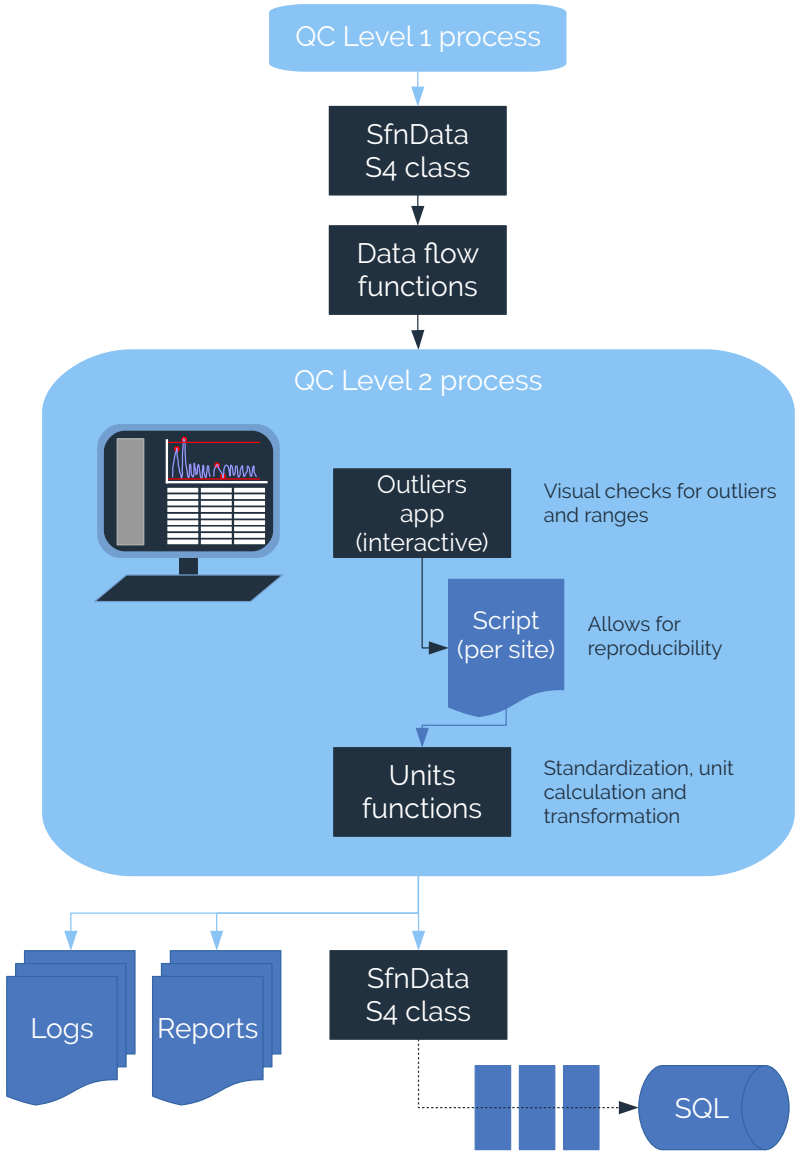
SfnData class:



SAPFLUXNET Data Flow



SAPFLUXNET Data Flow



SAPFLUXNET Status File

FOO_BAR_BAZ_status.yaml

```
QC:
  DONE: yes
  DATE: '2017-01-10'

LVL1:
  STORED: yes
  DATE: '2017-01-10'
  TO_LVL2: FREEZE/READY/DONE

LVL2:
  STORED: no
  DATE: ~
```

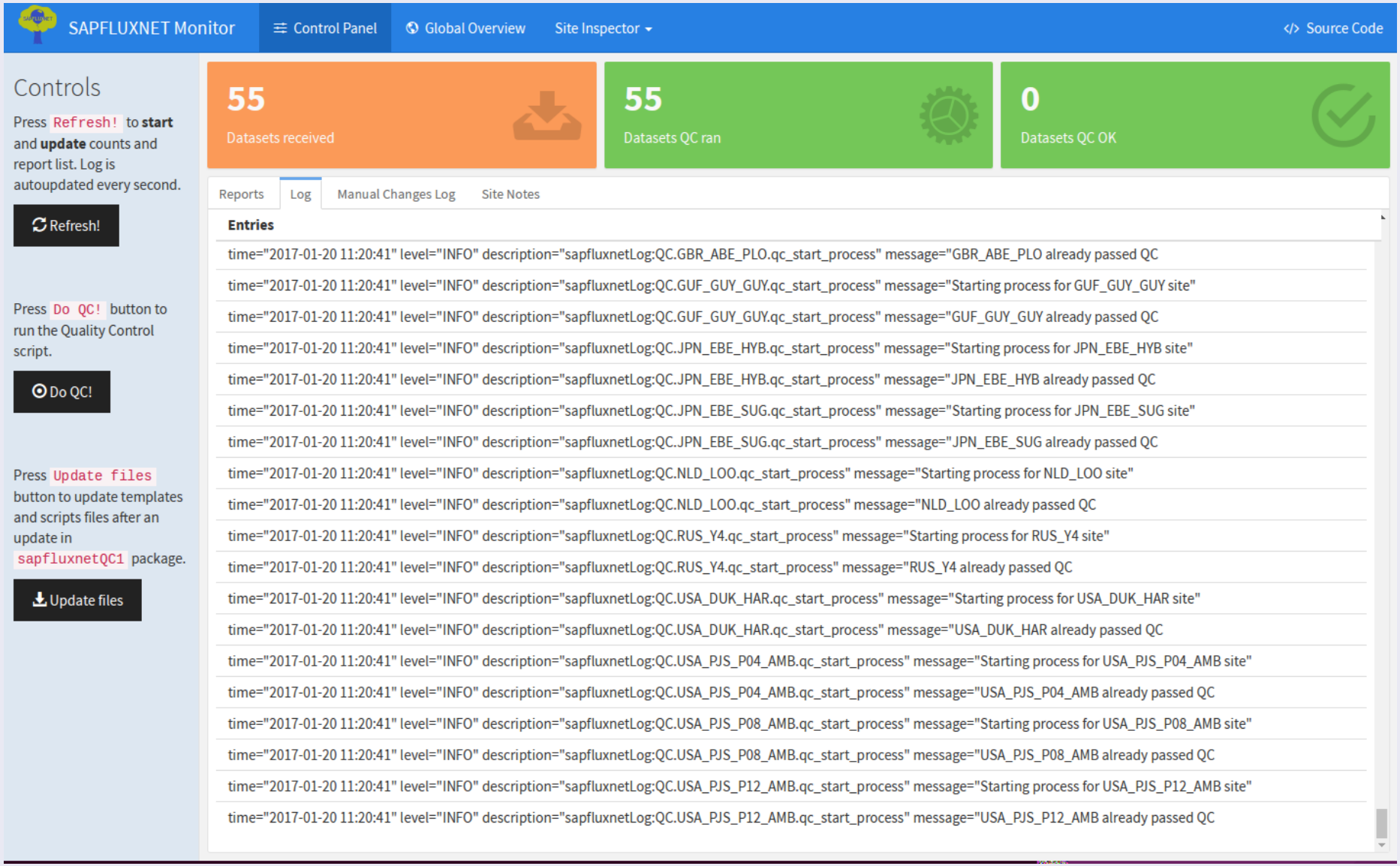
How to track site status?

Status file in yaml format:

- Known standard (i.e. used as headers in Rmarkdown)
- Easy to import/export in R
- Fast method to know the status

SAPFLUXNET Apps

Control center



SAPFLUXNET Apps

Progress dashboard

SAPFLUXNET Progress Report

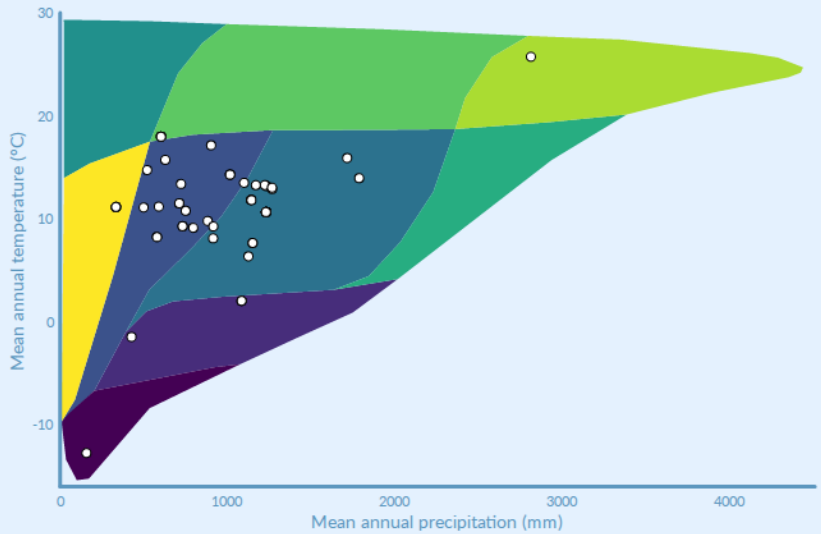
SAPFLUXNET Project Wiki

Start | Sites | Biomes | Methods | Species | Contributors

55
Sites

7
Biomes

Explore the biomes we have already. Hover the points to see the site code, and click in the table to view only the desired biome/s



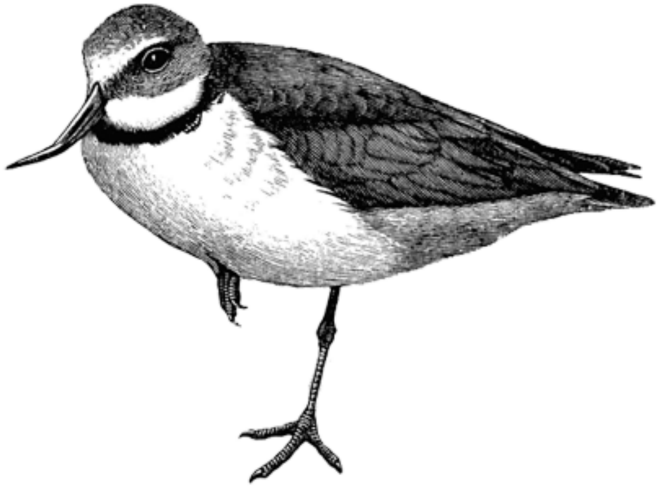
Biomes
- Boreal forest
- Mediterranean
- Subtropical desert
- Temperate forest
- Temperate grassland desert
- Temperate rain forest
- Tropical forest savanna
- Tropical rain forest
- Tundra

	Biome	Number of sites
1	Boreal forest	4
2	Mediterranean	19
3	Temperate forest	23
4	Temperate grassland desert	3
5	Tropical forest savanna	4
6	Tropical rain forest	1
7	Tundra	1

Showing 1 to 7 of 7 entries

Made by the SAPFLUXNET Team
with shiny and shinydashboard

Because contributing to SAPFLUXNET is cool



Contribute to SAPFLUXNET!!

Lets make the sap flow!

O RLY?

The SAPFLUXNET Team