



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

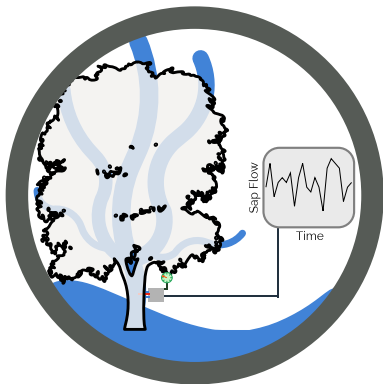
XIV MEDECOS & XIII AEET meeting

Ecoinformatics: data science brings new avenues for ecology
Symposium

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



Centre of Ecological Research and Forestry Applications



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

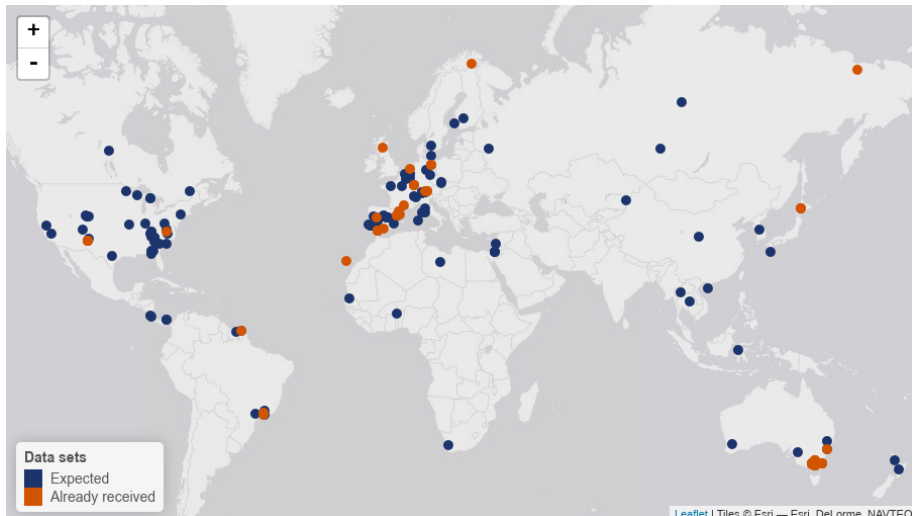
Different **thermodynamic methods** to determine sap flow using heat as a tracer sap movement

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 physiological factors driving tree- and stand-level transpiration



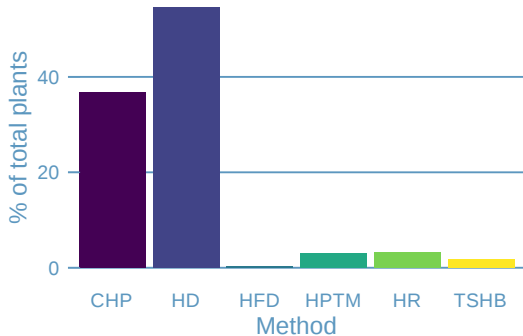


- ▶ Stem or whole-plant level
- ▶ Field conditions
- ▶ Sub-daily intervals
- ▶ Environmental data available (RH, Ta, PAR...)
- ▶ Abundant metadata (site, stand, plant, species and environmental)



High data complexity:

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

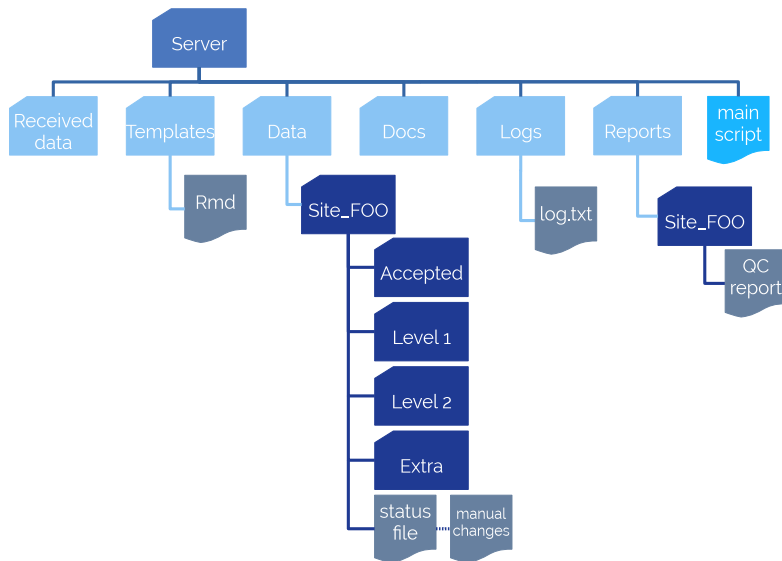


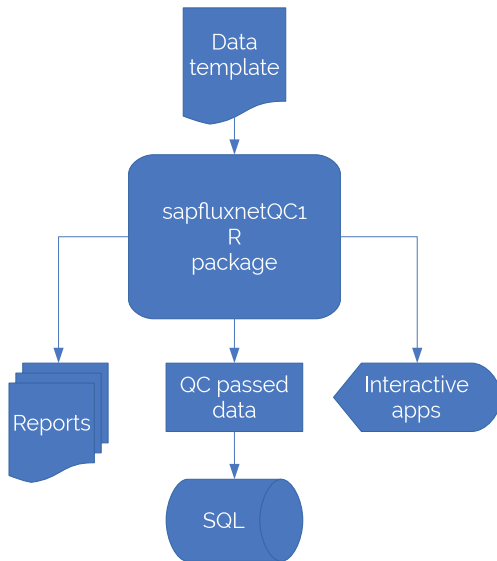


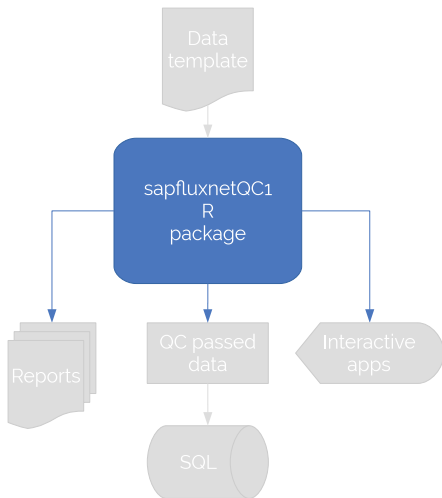
How to assay quality and store data?

We need **semi-automatic**, **reproducible** and **robust** checks to ensure the quality of the submitted 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

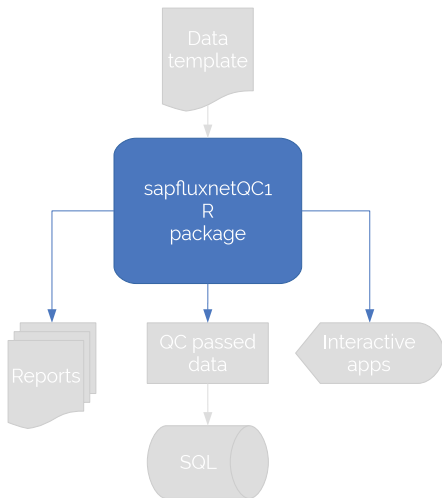
SAPFLUXNET Server Structure







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



Benefits of R as development environment

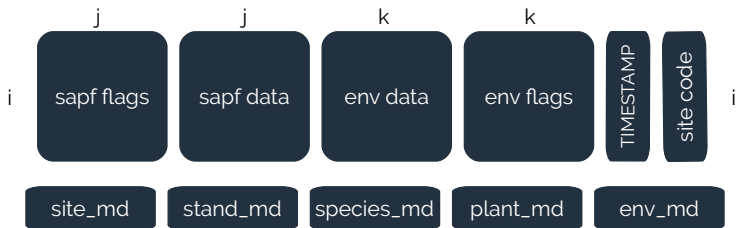
- ▶ Open
- ▶ Reproducible
- ▶ Easy maintenance and update
- ▶ Easy integration with web and SQL technologies

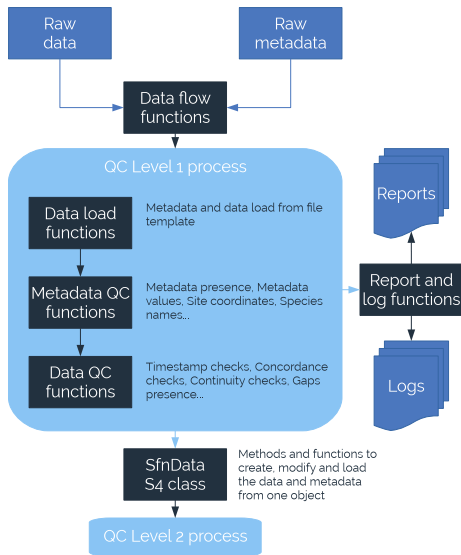


S4 classes:

- ▶ 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:

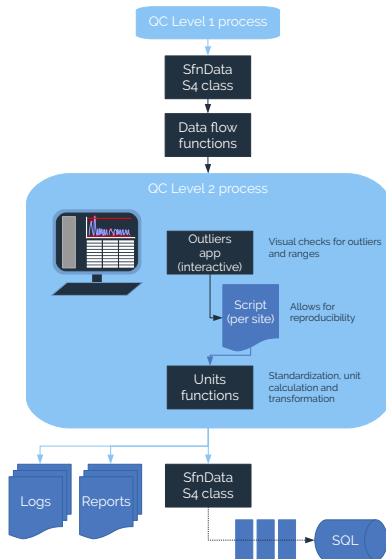




QC Level 1

General metadata and data quality checks:

- ▶ Presence/Absence of metadata and data variables
- ▶ Metadata values correctness (i.e coordinates, species names...)
- ▶ Sapflow and environmental data correctness (format, timezone, gaps, continuity...)
- ▶ Uniformization and unit transformations of data (solar time, sapflow at different levels...)



QC Level 2

Specific Data quality checks:

- ▶ Robust outlier detection
- ▶ Range checks
- ▶ Unit standarization and transformation



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



Control center

SAPFLUXNET Monitor

≡ Control Panel
 [Global Overview](#)
[Site Inspector](#)

[Source Code](#)

Controls

Press **Refresh!** to start and **update** counts and report list. Log is autoupdated every second.

Refresh!

Press **Do QC!** button to run the Quality Control script.

Do QC!

Press **Update files** button to update templates and scripts files after an update in **sapfluxnetQC1** package.

Update files

55

Datasets received

55

Datasets QC ran

0

Datasets QC OK

Reports

Log

Manual Changes Log

Site Notes

Entries

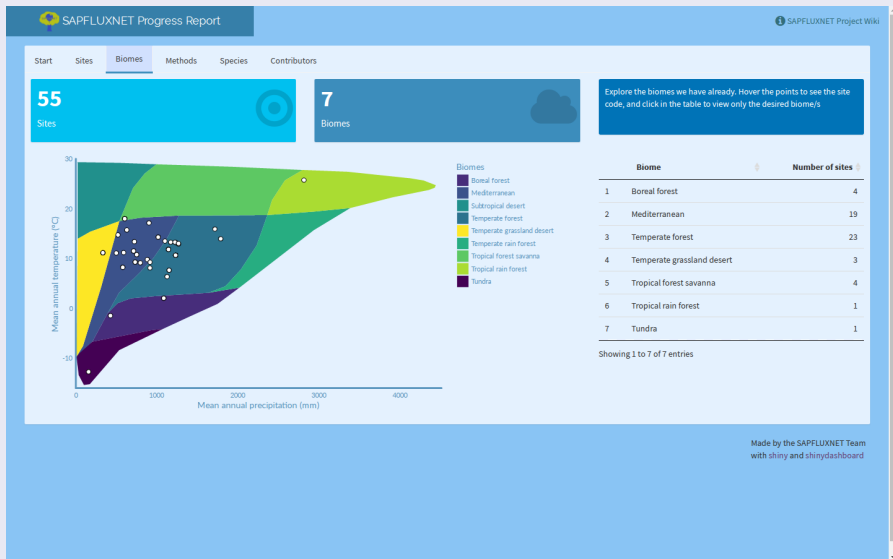
```

time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.GBR_ABE_PLO.qc_start_process" message="GBR_ABE_PLO already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.GUF_GUY_GUY.qc_start_process" message="Starting process for GUF_GUY_GUY site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.GUF_GUY_GUY.qc_start_process" message="GUF_GUY_GUY already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.JPN_EBE_HYB.qc_start_process" message="Starting process for JPN_EBE_HYB site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.JPN_EBE_HYB.qc_start_process" message="JPN_EBE_HYB already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.JPN_EBE_SUG.qc_start_process" message="Starting process for JPN_EBE_SUG site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.JPN_EBE_SUG.qc_start_process" message="JPN_EBE_SUG already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.NLD_LOO.qc_start_process" message="Starting process for NLD_LOO site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.NLD_LOO.qc_start_process" message="NLD_LOO already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.RUS_Y4.qc_start_process" message="Starting process for RUS_Y4 site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.RUS_Y4.qc_start_process" message="RUS_Y4 already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.USA_DUK_HAR.qc_start_process" message="Starting process for USA_DUK_HAR site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.USA_DUK_HAR.qc_start_process" message="USA_DUK_HAR already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.USA_PJS_P04_AMB.qc_start_process" message="Starting process for USA_PJS_P04_AMB site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.USA_PJS_P04_AMB.qc_start_process" message="USA_PJS_P04_AMB already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.USA_PJS_P08_AMB.qc_start_process" message="Starting process for USA_PJS_P08_AMB site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.USA_PJS_P08_AMB.qc_start_process" message="USA_PJS_P08_AMB already passed QC
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.USA_PJS_P12_AMB.qc_start_process" message="Starting process for USA_PJS_P12_AMB site"
time="2017-01-20 11:20:41" level="INFO" description="sapfluxnetLog:QC.USA_PJS_P12_AMB.qc_start_process" message="USA_PJS_P12_AMB already passed QC

```

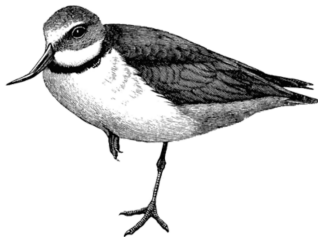


Progress dashboard





Because contributing to SAPFLUXNET is cool



Contribute to
SAPFLUXNET!!

Lets make the sap flow!

O RLY?

The SAPFLUXNET Team