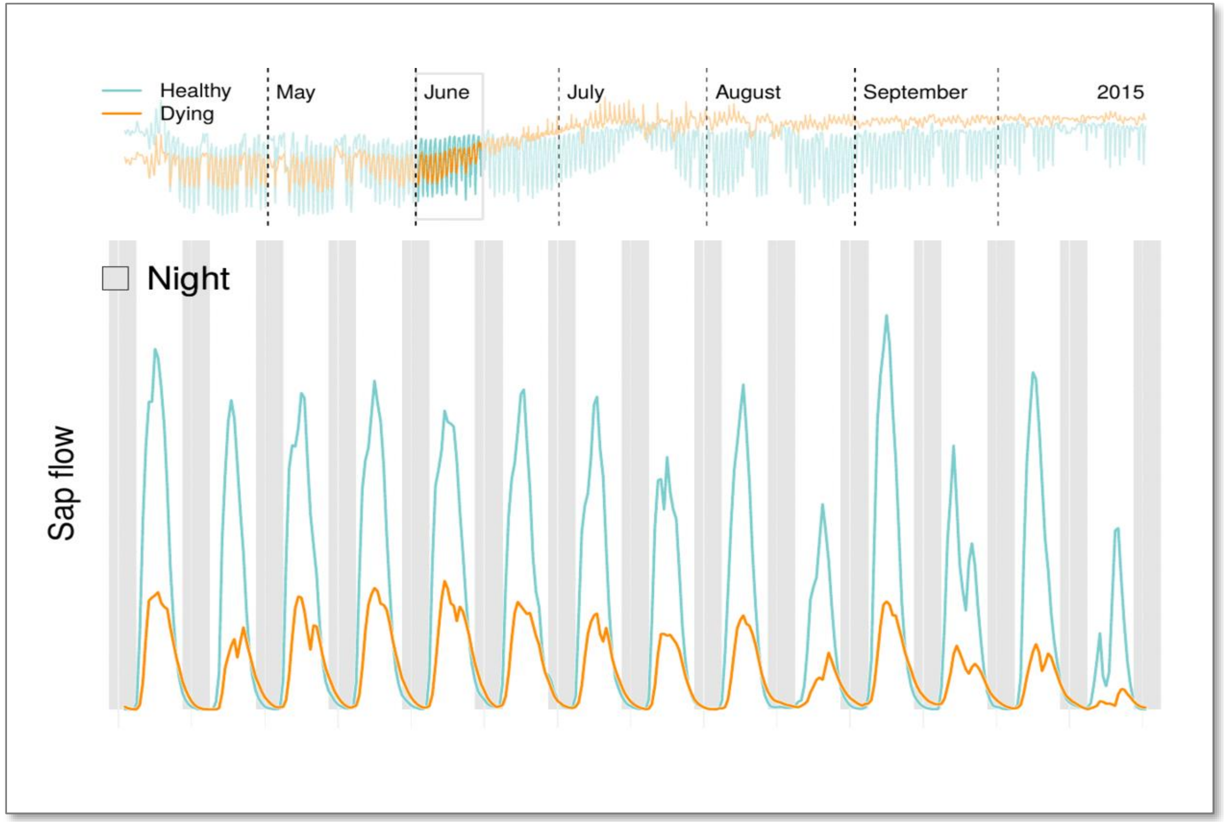
Time series data processing

Goal? *Make methods for data processing in ecological research more accessible, save time and standardize workflows.*



University of Helsinki
Environmental Data Science

Richard L. Peters

LIÈGE université Gembloux Agro-Bio Tech
GHENT UNIVERSITY
WSL

Alexander G. Hurley

GFZ Helmholtz Centre Potsdam

Relevance? *Ecological research is becoming increasingly data-rich!*

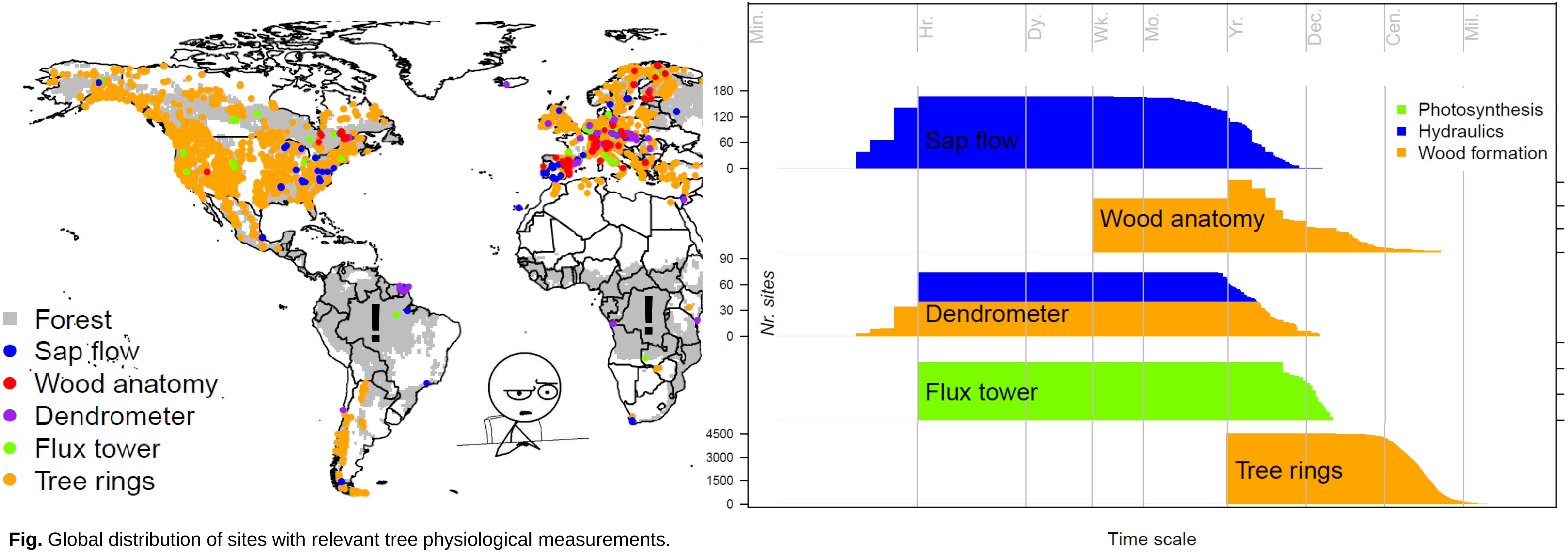
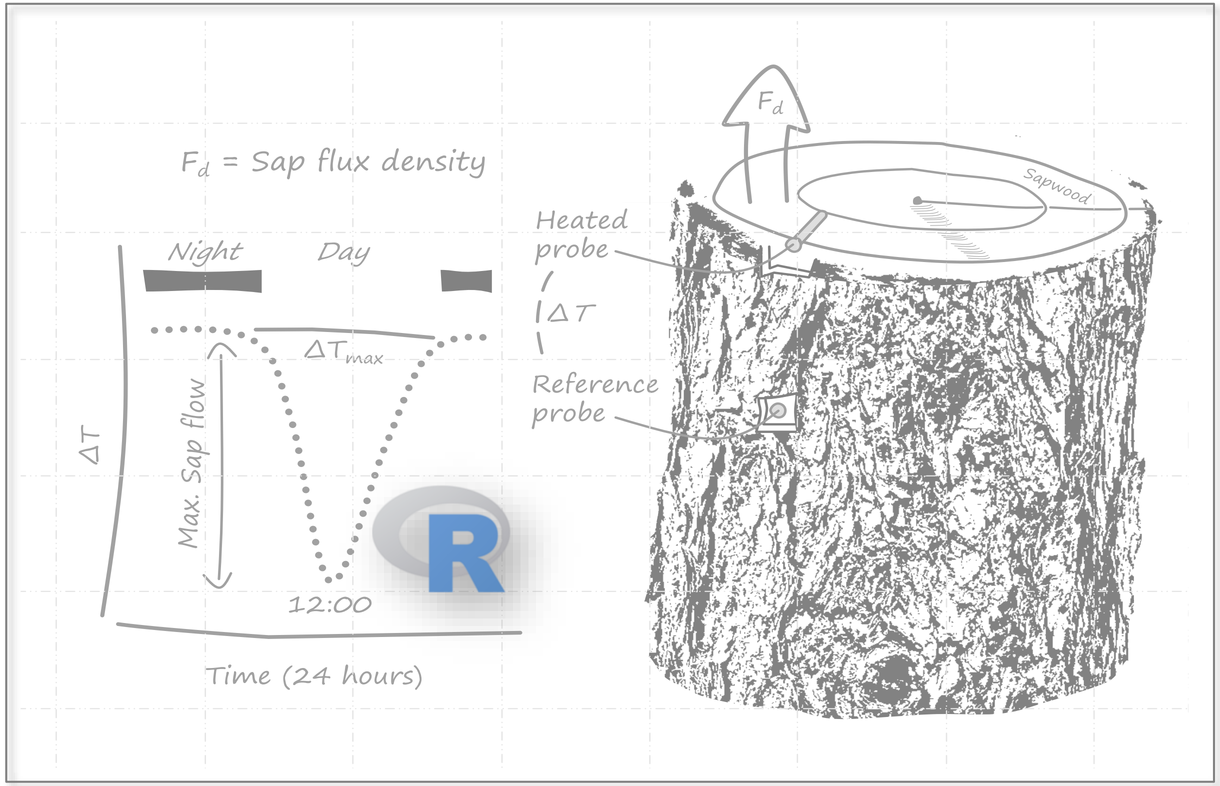


Fig. Global distribution of sites with relevant tree physiological measurements.

Source: Peters (2018) doi: https://doi.org/10.5451/unibas-007085812

Goal? *Make methods for data processing in ecological research more accessible, save time and standardize workflows.*

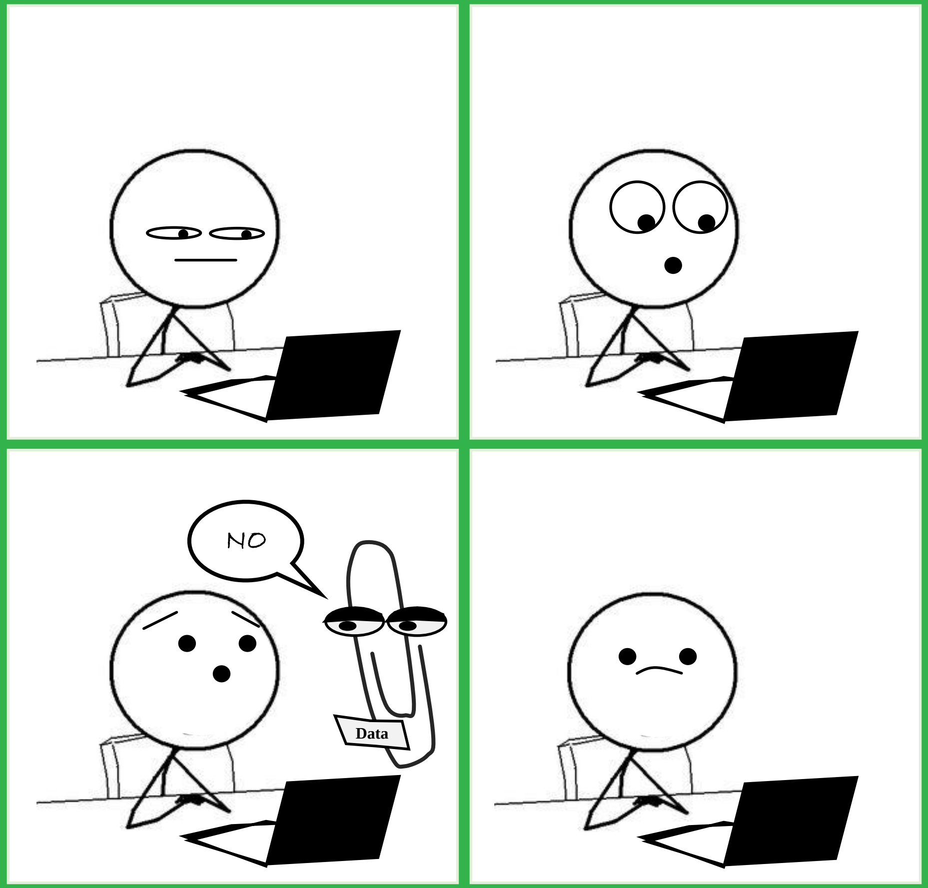


University of Helsinki
Environmental Data Science

Course
Introduction
Typical data issues
Programming solutions
Monitoring examples
(Sap flow + Dendrometers)

Assignments
01_datacleanr
02a_treenetproc
02b_trex

Typical data issues



Typical data issues

Time series ≈ *A collection of values obtained over time*

Fig. Examples of relevant reference for performing time series analyses, with a focus on its application.

See also: Zuur *et al.* (2009) doi: https://doi.org/10.1007/978-0-387-87458-6

Focus! *Obtaining "clean" data usable for data analysis*

Properties of time-series data:

- Large quantities;
- ±Regular collection intervals;
- Often requires data aggregation.

Processing such data is challenging:

- Data formatting issues;
- Timestamp and interval issues;
- Outlier and sensor failure issues.

Typical data issues

Data formatting issues *Three typical examples*

Long format

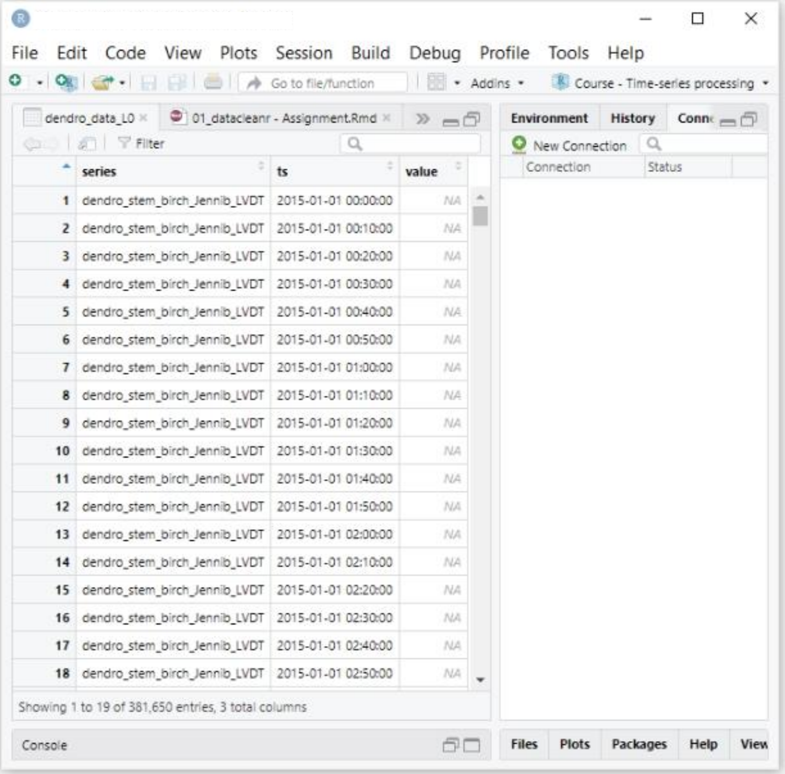
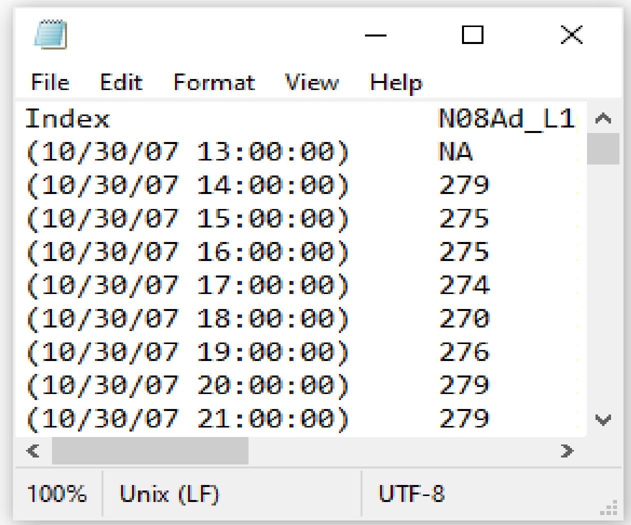
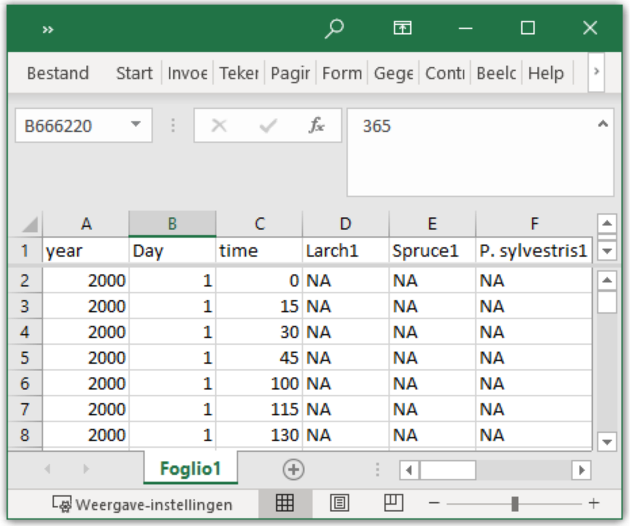


Fig. Sites with dendrometer and sap flow data.

Wide format



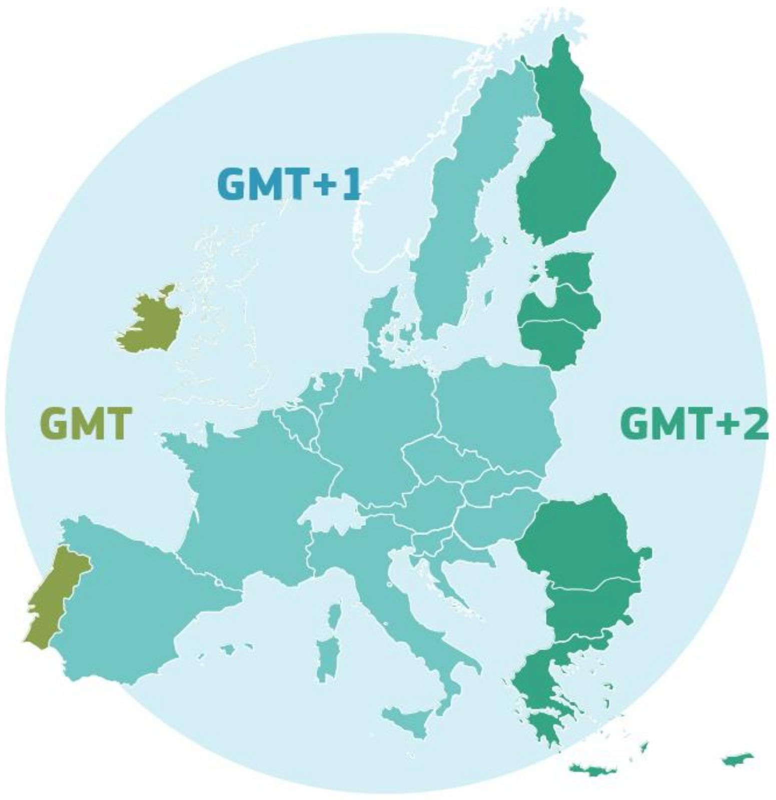
Timestamp



Day of year

Typical data issues

Timestamp and interval issues *Time zone and the elephant in the room (daylight saving time…)*



Source: ec.europa.eu/transport/themes/summertime_en

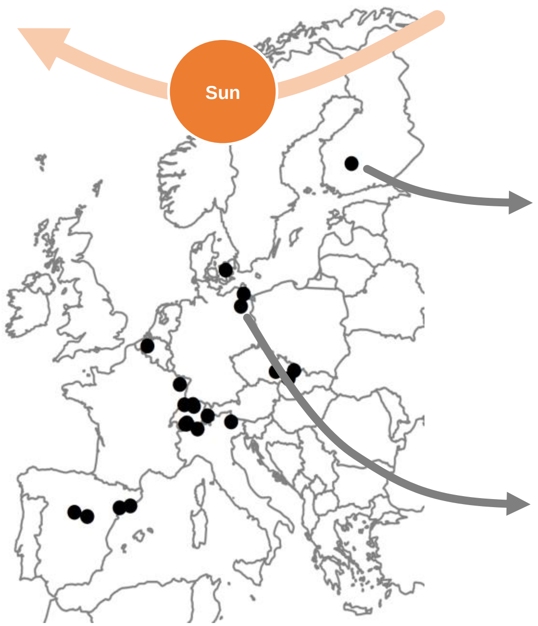
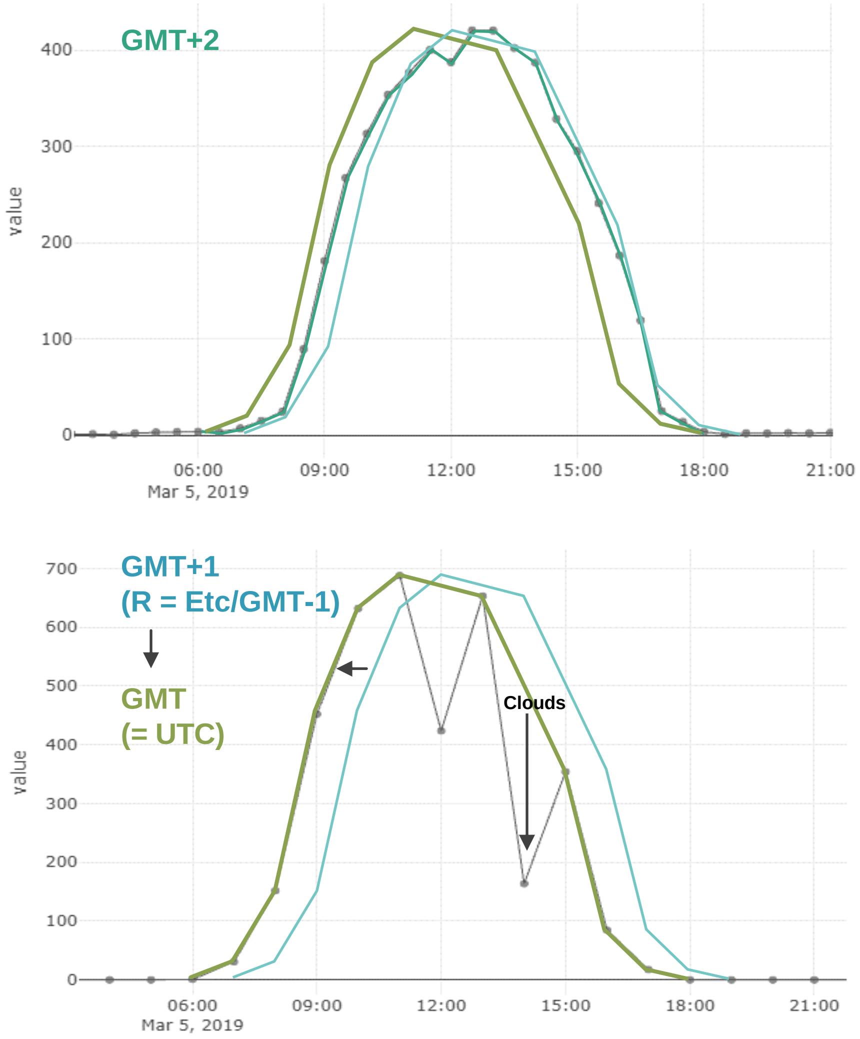
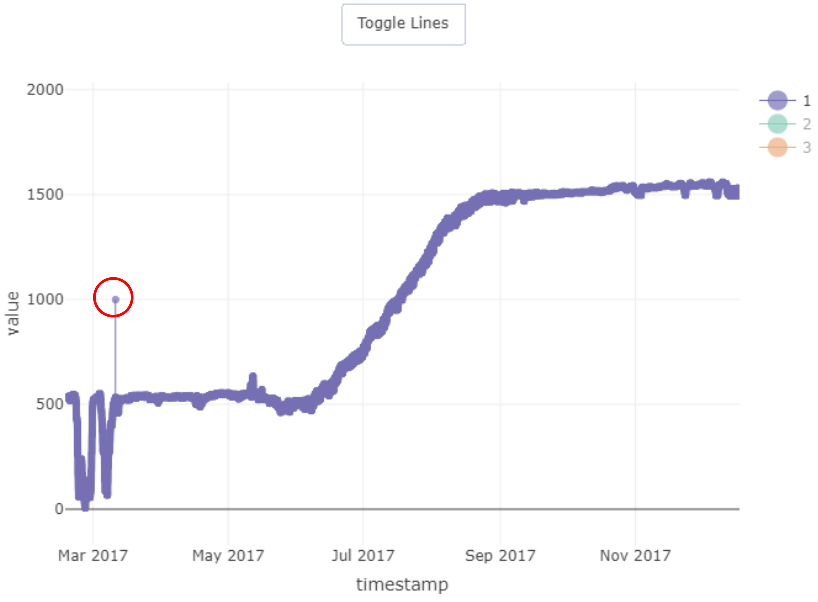


Fig. Sites with global irradiance (W/m²) data.

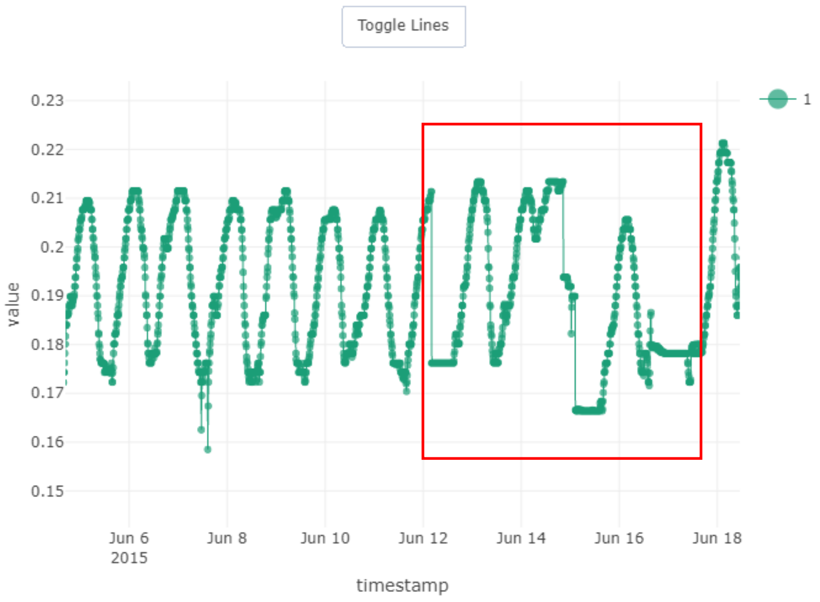


Outlier and sensor failure issues *Removing data should always be done with care!*

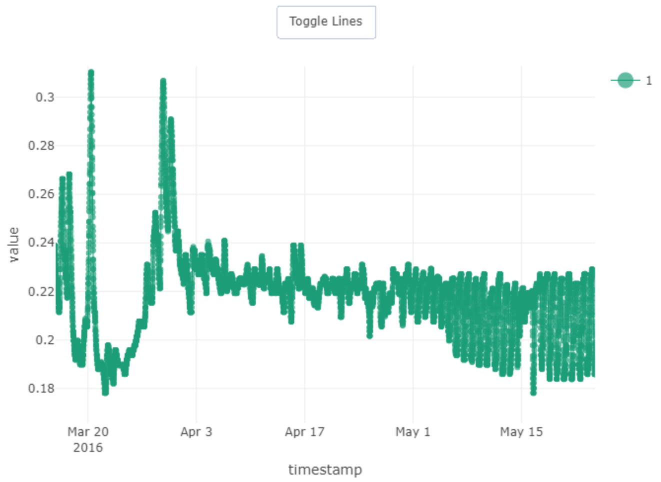
Outliers (dendrometer data)

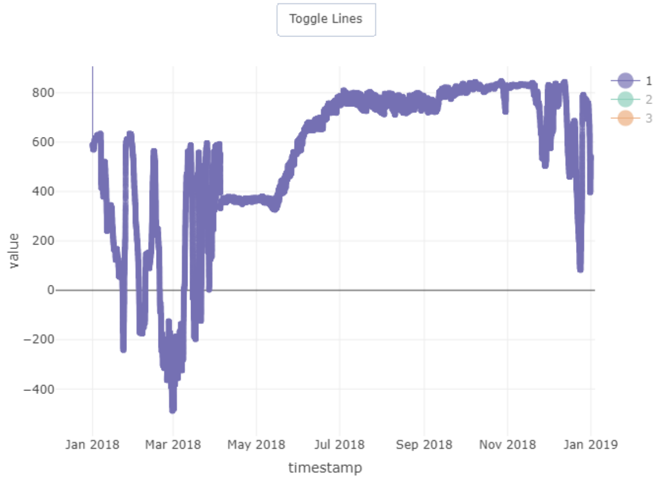


Sensor failure (sap flow data)

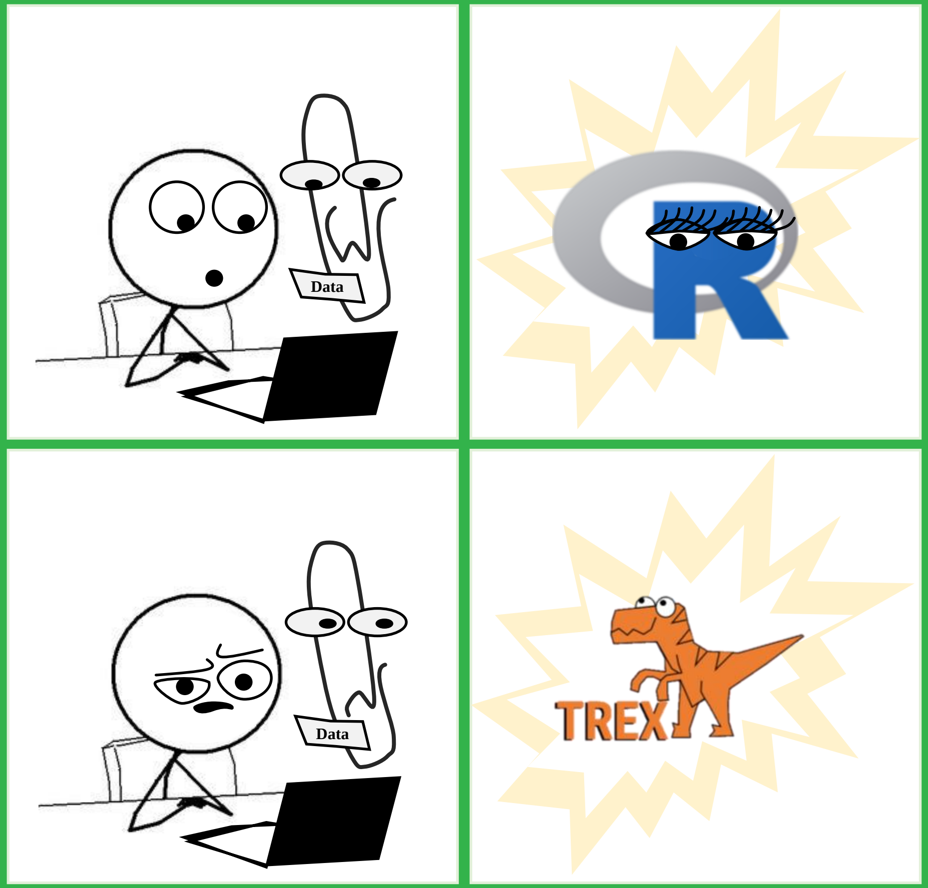


What about spring data?





Programming solutions



Programming solutions

Available software *Excellent software is available to work with time-series data*

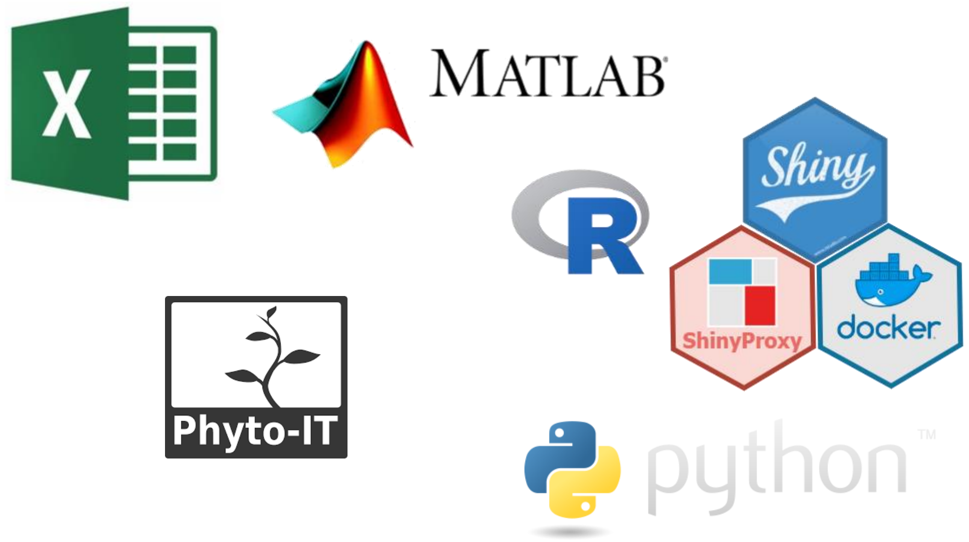


Fig. Examples of software available for performing time-series data processing. Within this course we will focus on using R, RStudio, R Shiny apps and specific CRAN packages.

Focus! *R is extensively used in Environmental Sciences*

Packages are key:

- zoo;
- lubridate;
- datacleanr;

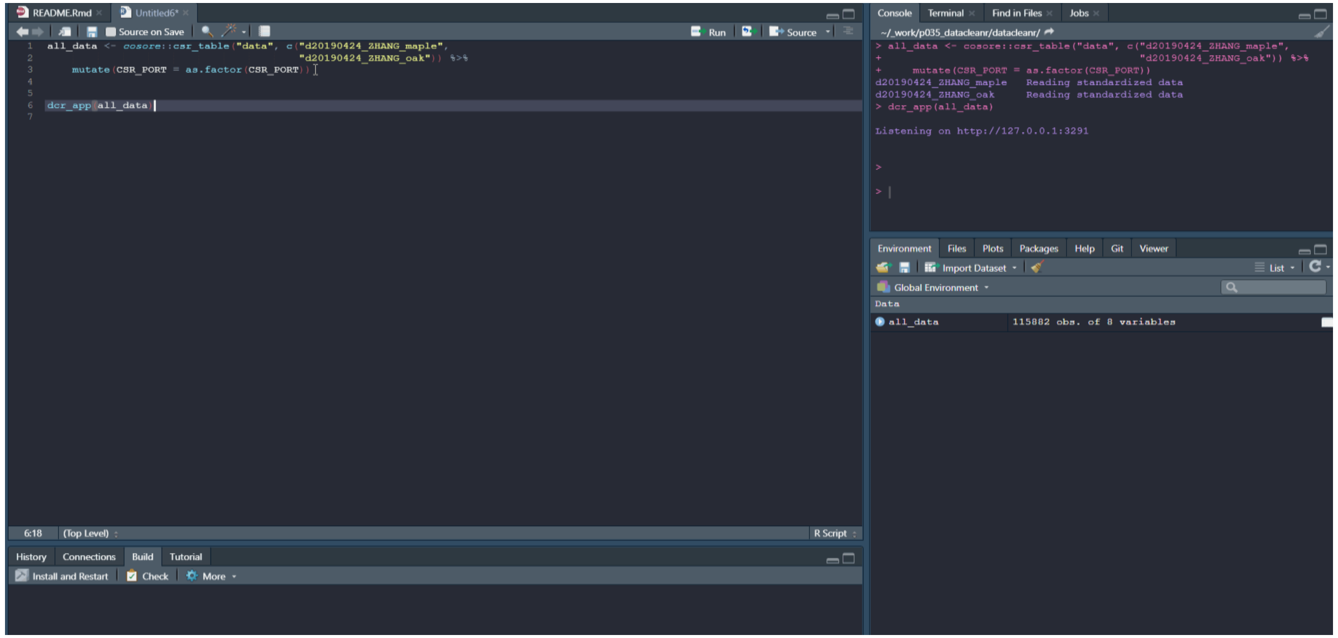
```
# additional installs for later use
packages <- (c("remotes", "dplyr", "forcats","readxl", "lubridate"))
install.packages(setdiff(packages,
                         rownames(installed.packages())))
remotes::install_github("the-hull/datacleanr")

library(datacleanr)
library(dplyr)
```

- and many more…

See also: www.datacamp.com/community/tutorials/time-series-r

A novel R-based package *A flexible and efficient tool for interactive data cleaning*

datacleanr *Removing data should finally be reproducible! – A.G. Hurley*



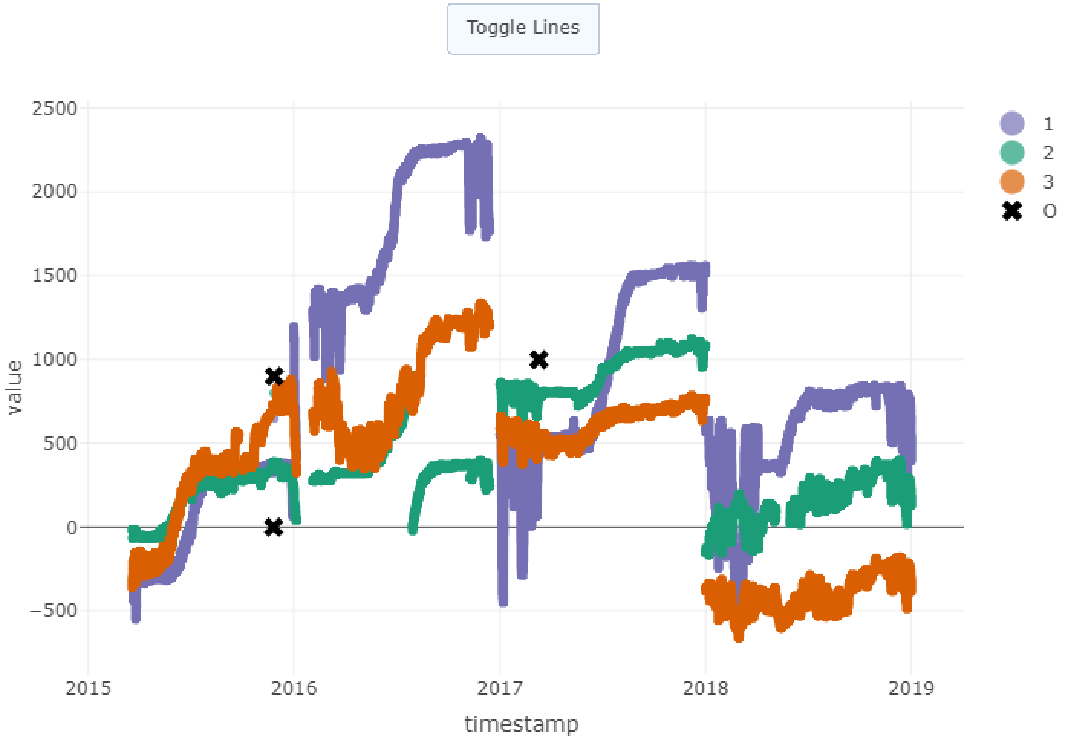
Properties:

- Uses R (links with other packages);
- Freely available (avoid license costs);
- Uses R shiny (interactive approach).

Structure of the tool:

- Set-up & overview;
- Filtering;
- Visual cleaning and annotating;
- Extract.

Extract *Reproducible recipe to cook up some fresh data*



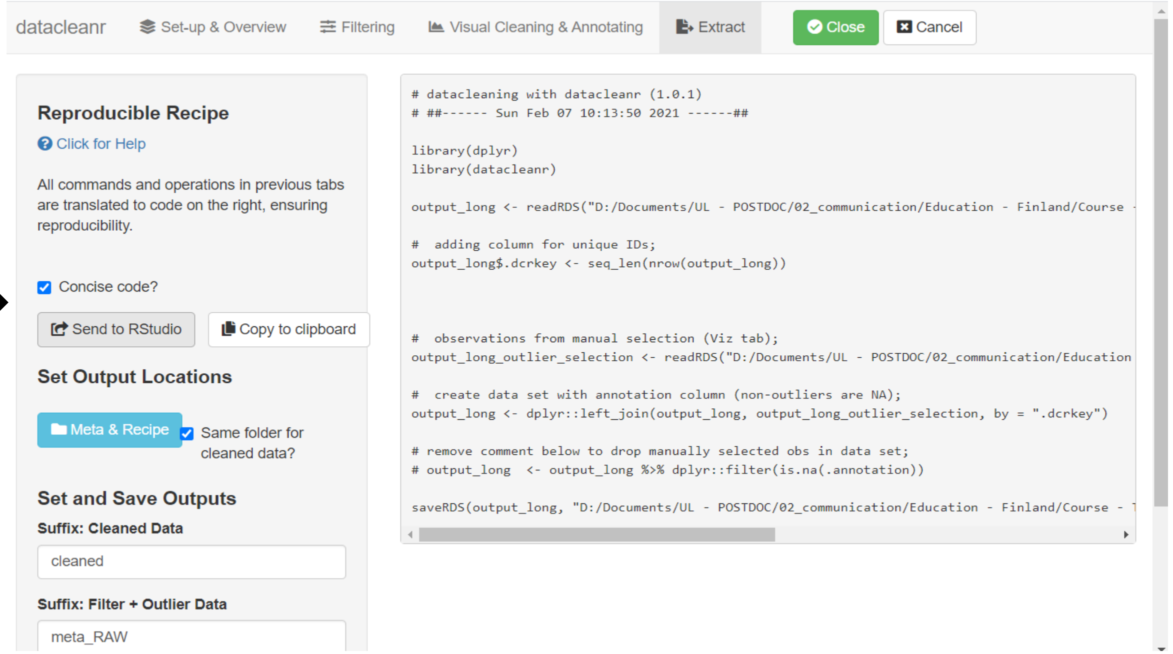


Fig. Example of dendrometer data recorded for three trees and its resulting datacleanr recipe.

Programming solutions

Versatile *Tested on multiple temporal- and spatial-specific data sets*

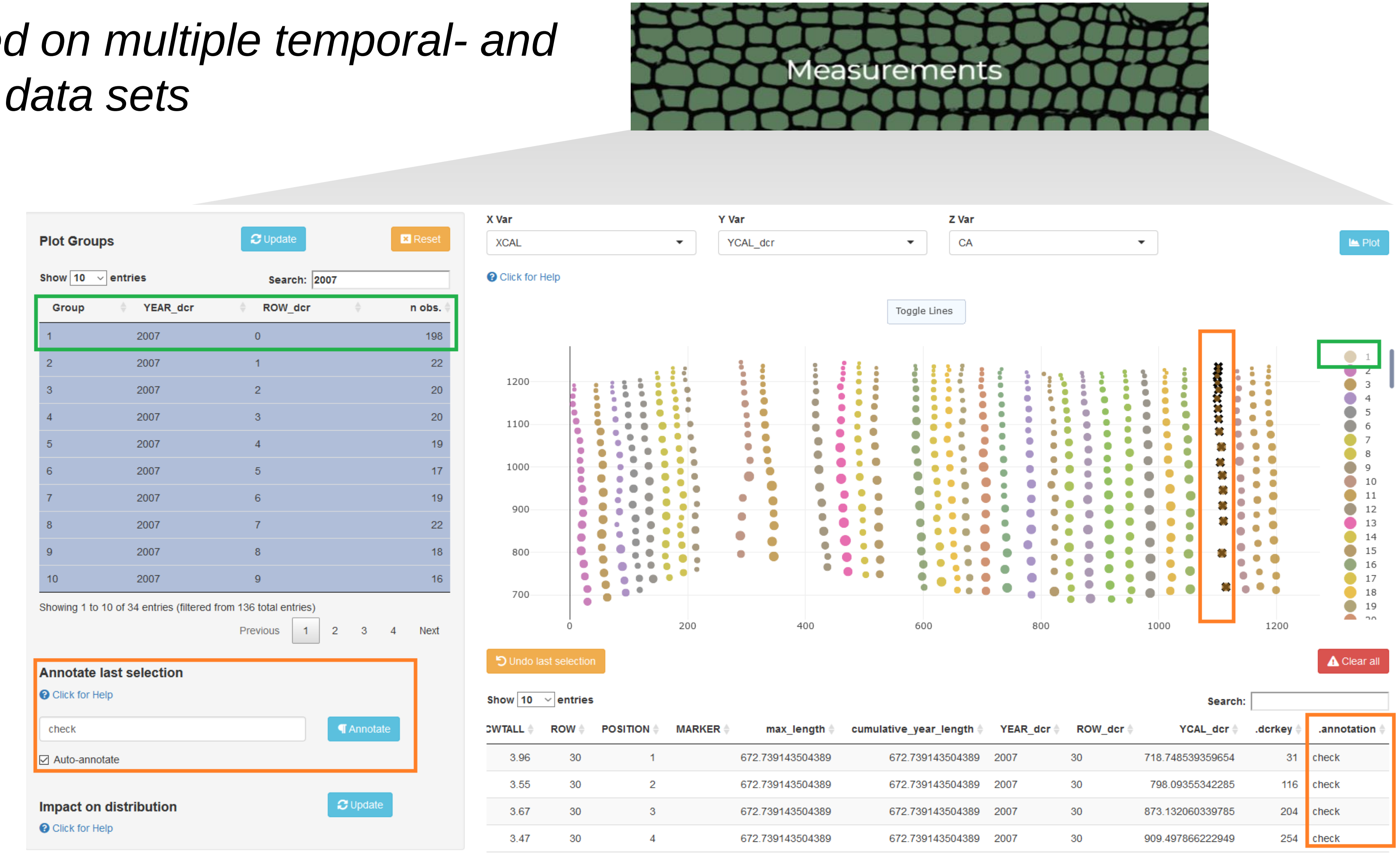
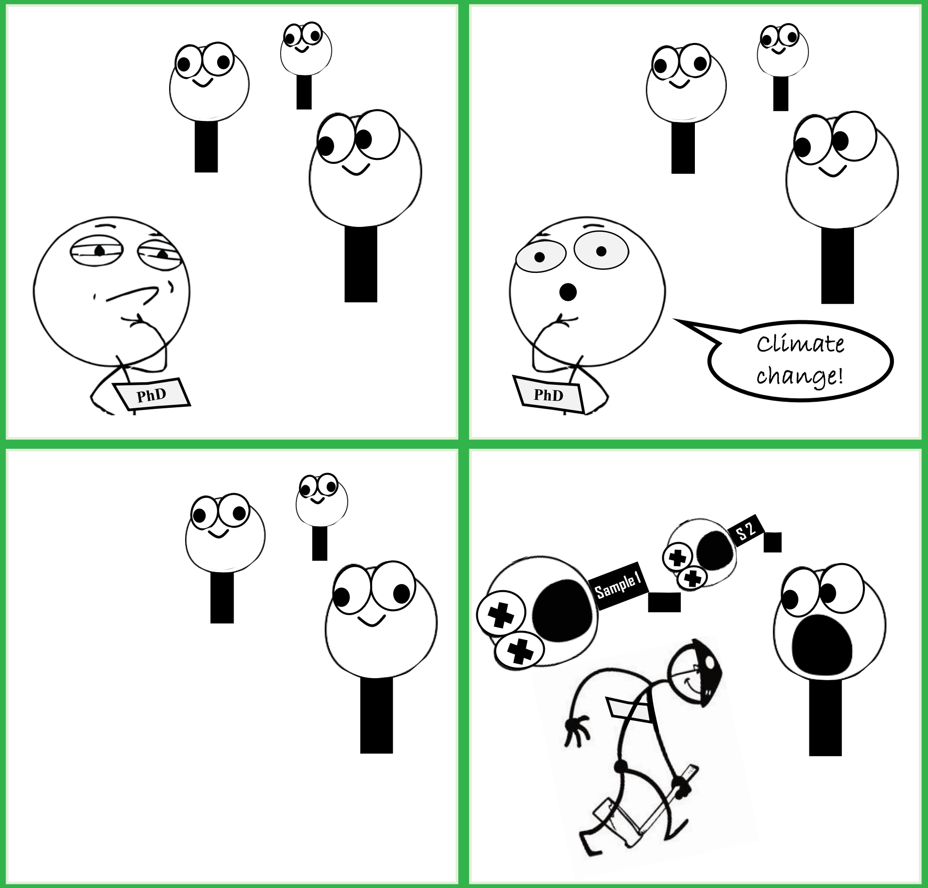


Fig. Example of wood anatomical data loaded into datacleanr. Each dot present a tracheid cell measured from a *Pinus cembra* tree (available within the RAPTOR R package).

Monitoring examples



Hyytiälä Forestry Field Station

Flux tower site in Finland

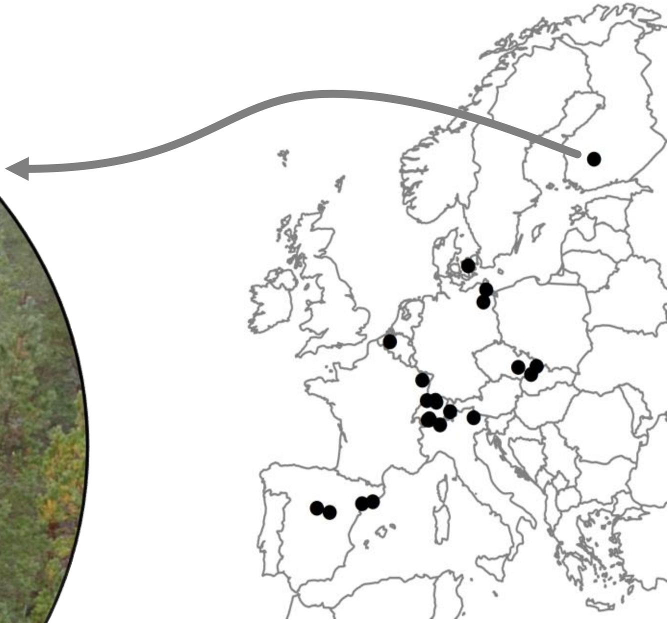
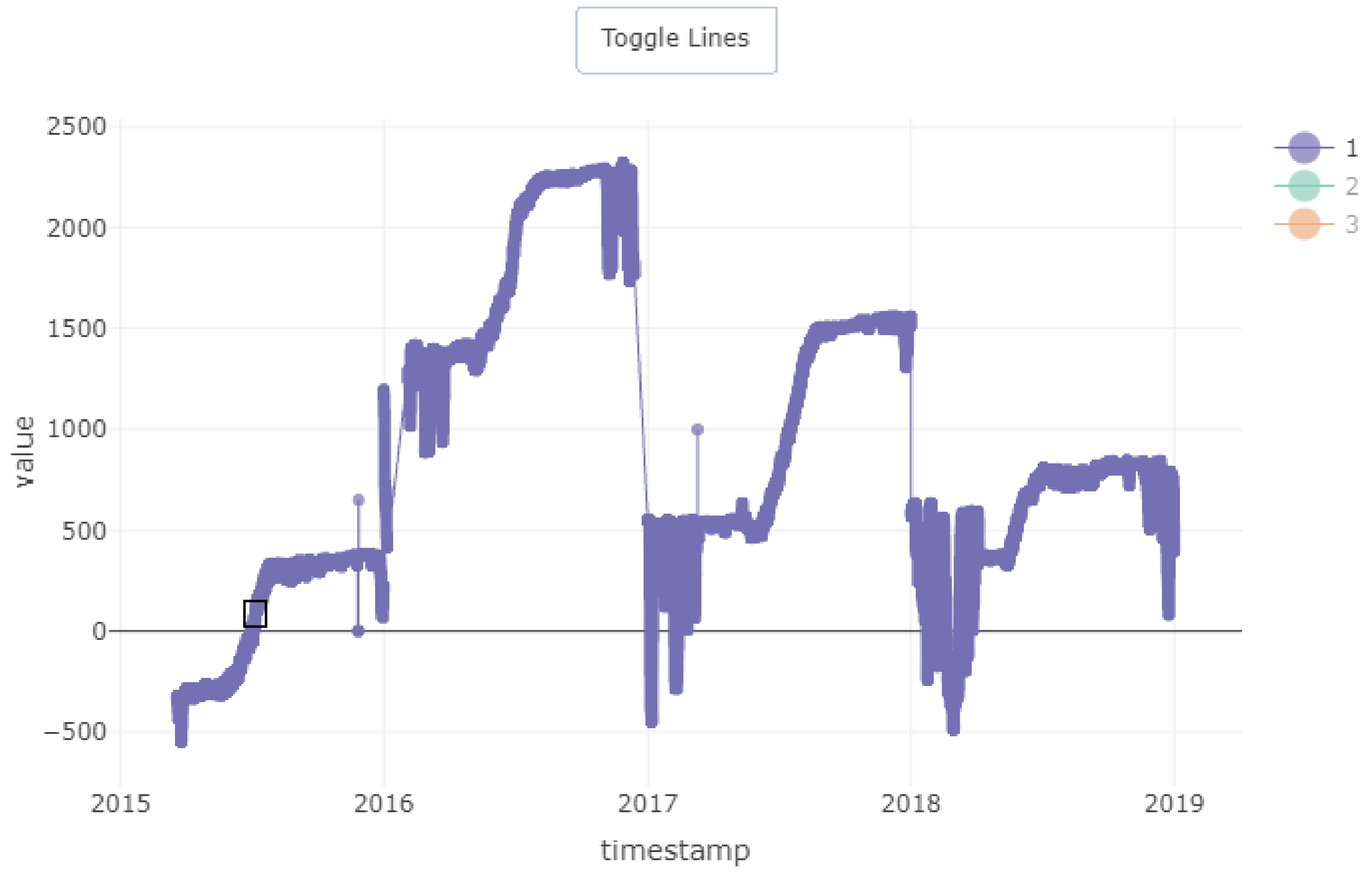
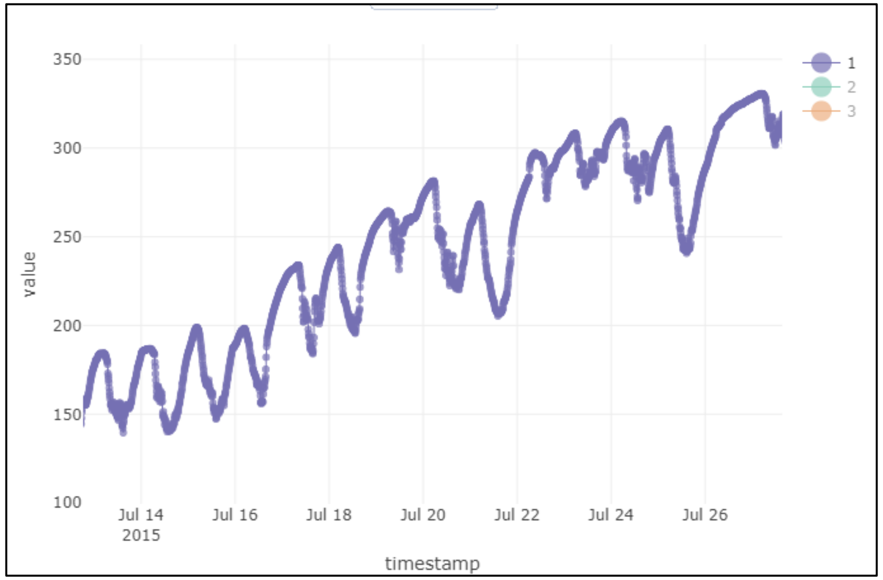


Fig. Sites with dendrometer and sap flow data.

Fig. Examples of an linear variable differential transformator (LVDT) dendrometer installed on a birch tree (top) and thermal dissipation probes (TDP) installed on a Scots pine (bottom), to measure stem radial change and sap flow, respectively.

Dendrometers *Data*





Dendrometers *Stem radius changes of trees provide information on growth and drought stress*

Fig. Examples of band and point dendrometers.

Usage *Isolating growth and reversible tree-water deficit shrinkage*

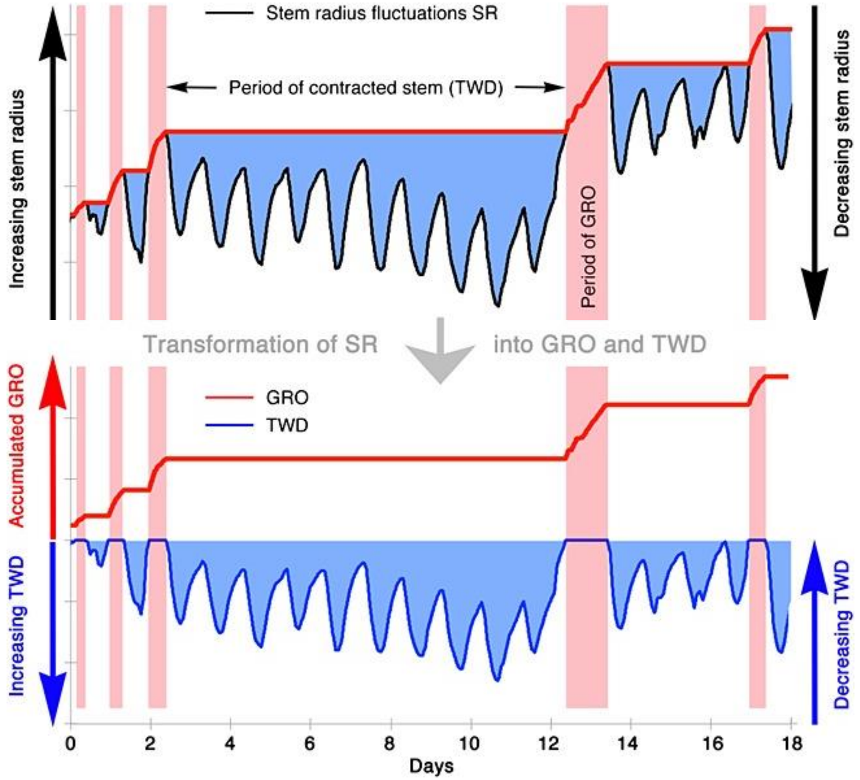
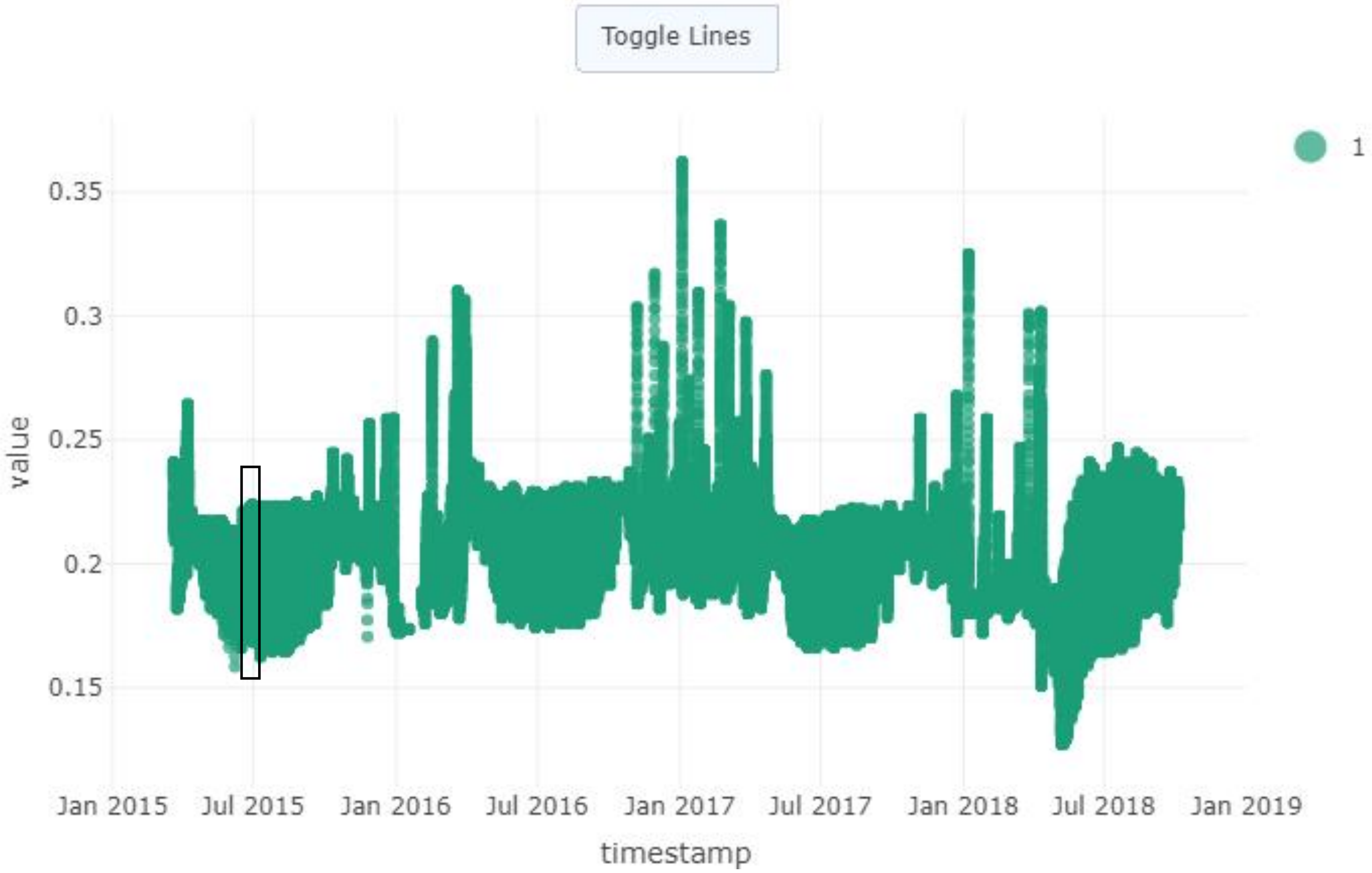
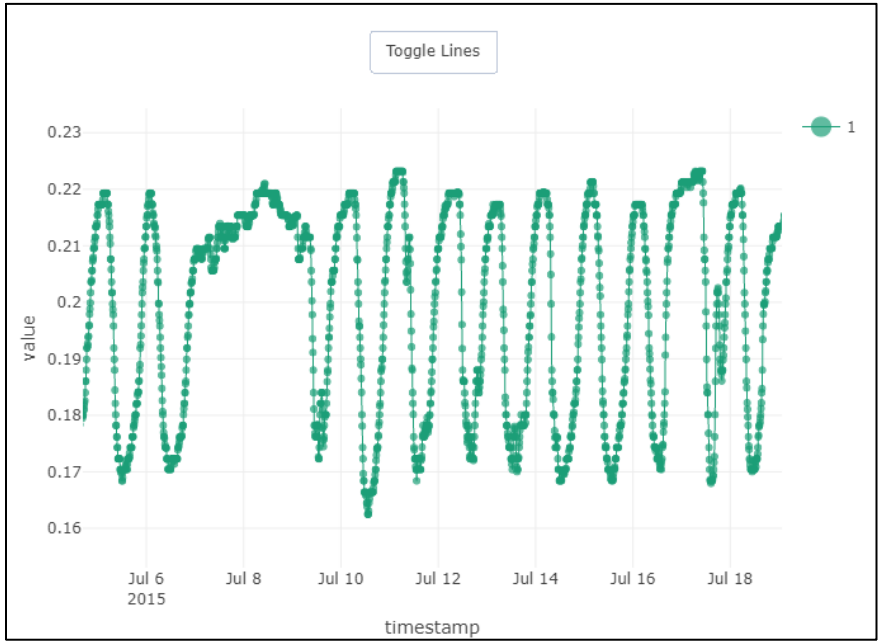


Fig. Example of a method to extract growth and drought stress indicators.

Source: Zweifel (2015) doi: https://doi.org/10.1111/pce.12613

Sap flow *Data*





Sap flow *Thermal dissipation probes to obtain water use*

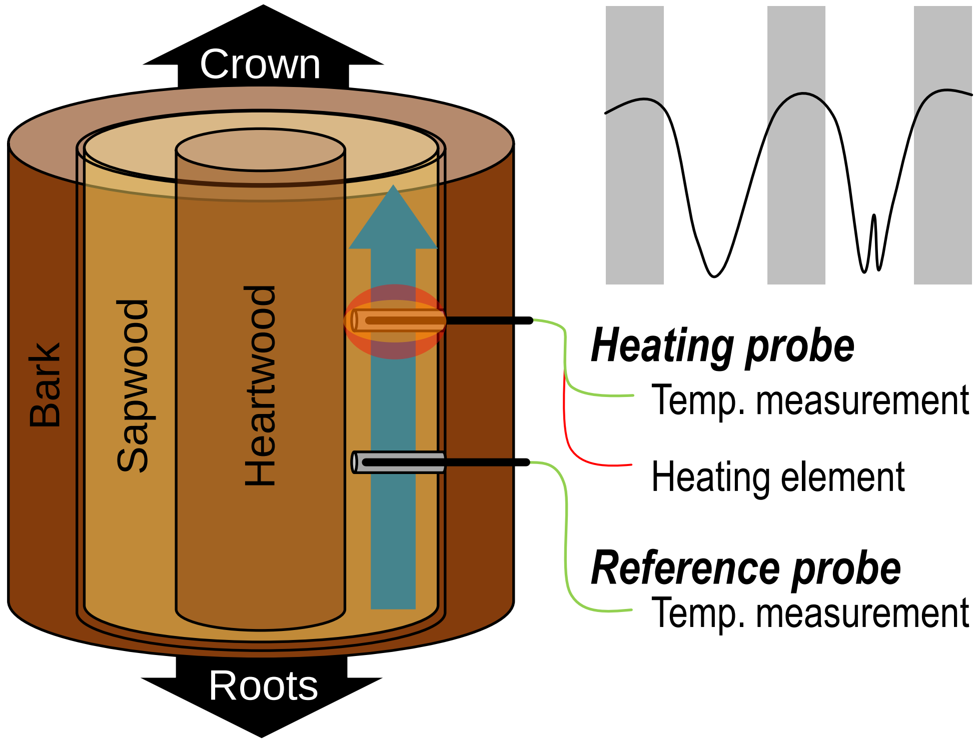


Fig. Schematics of thermal dissipation probes.

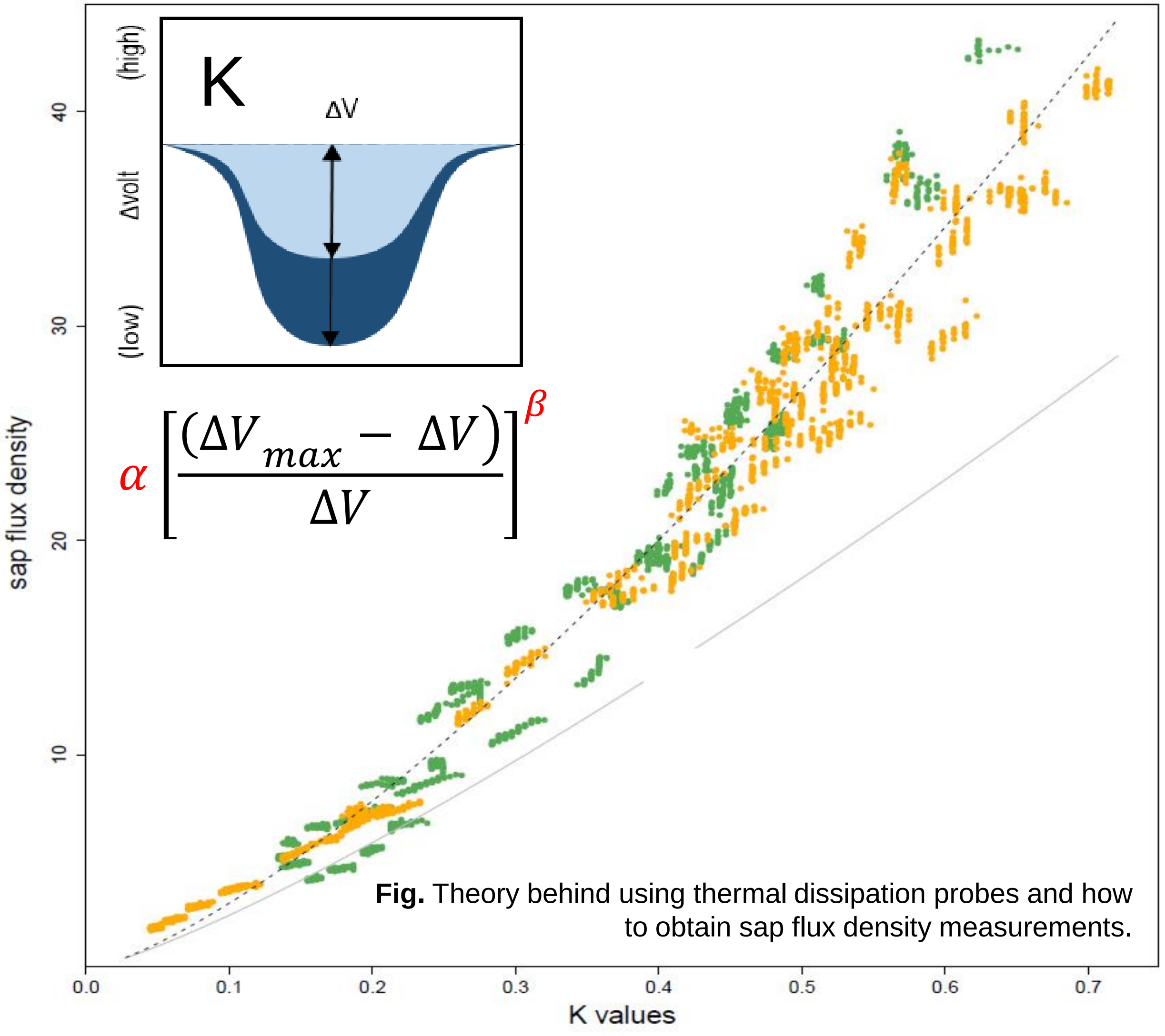
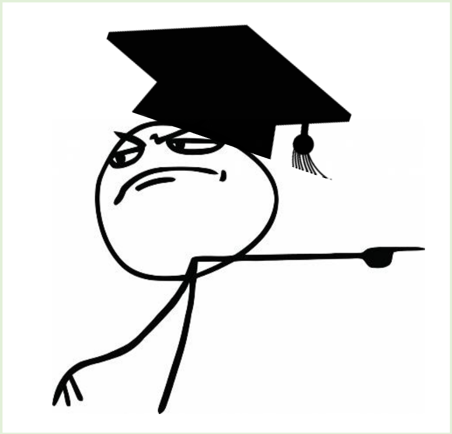
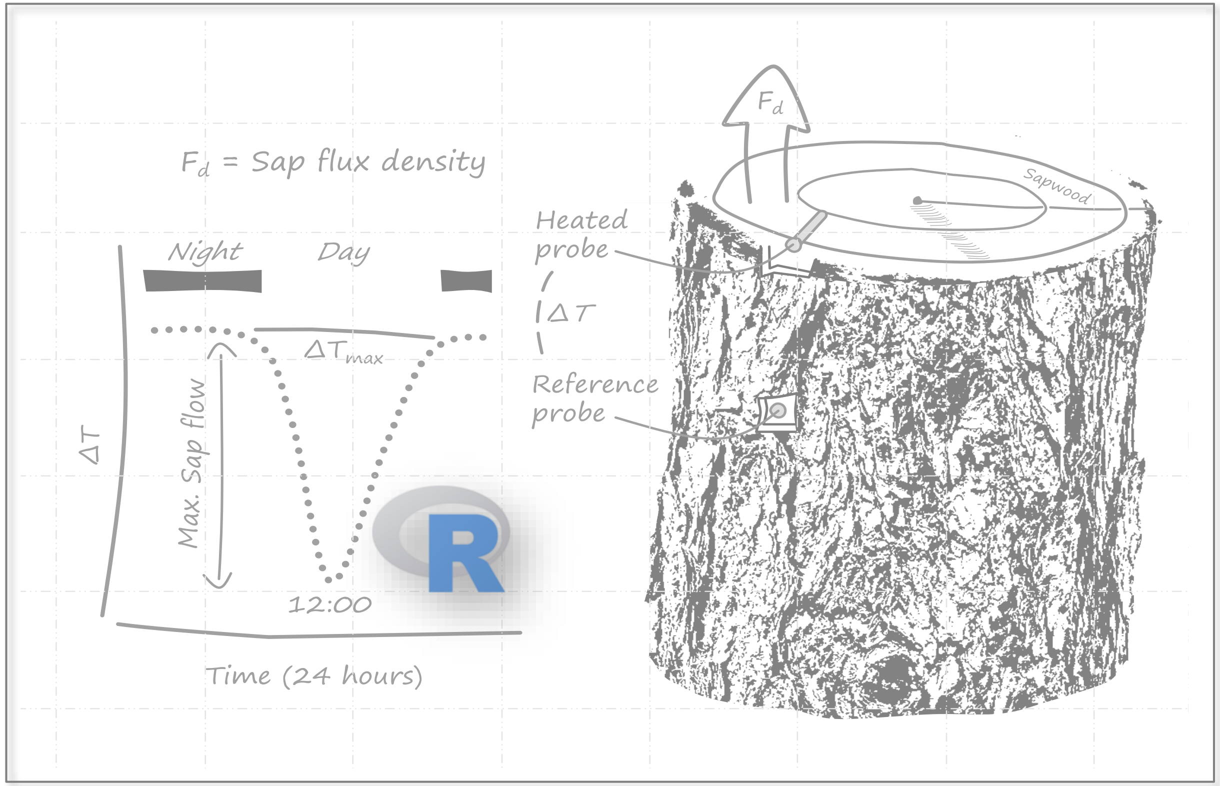


Fig. Theory behind using thermal dissipation probes and how to obtain sap flux density measurements.

Assignments



Goal? *Make methods for data processing in ecological research more accessible, save time and standardize workflows.*



University of Helsinki
Environmental Data Science

Course
Introduction
Typical data issues
Programming solutions
Monitoring examples
(Sap flow + Dendrometers)

Assignments
01_datacleanr
02a_treenetproc
02b_trex

Appendix

Time series data processing

Why these measurements? *Forests are anchor points in the global water and carbon cycle*

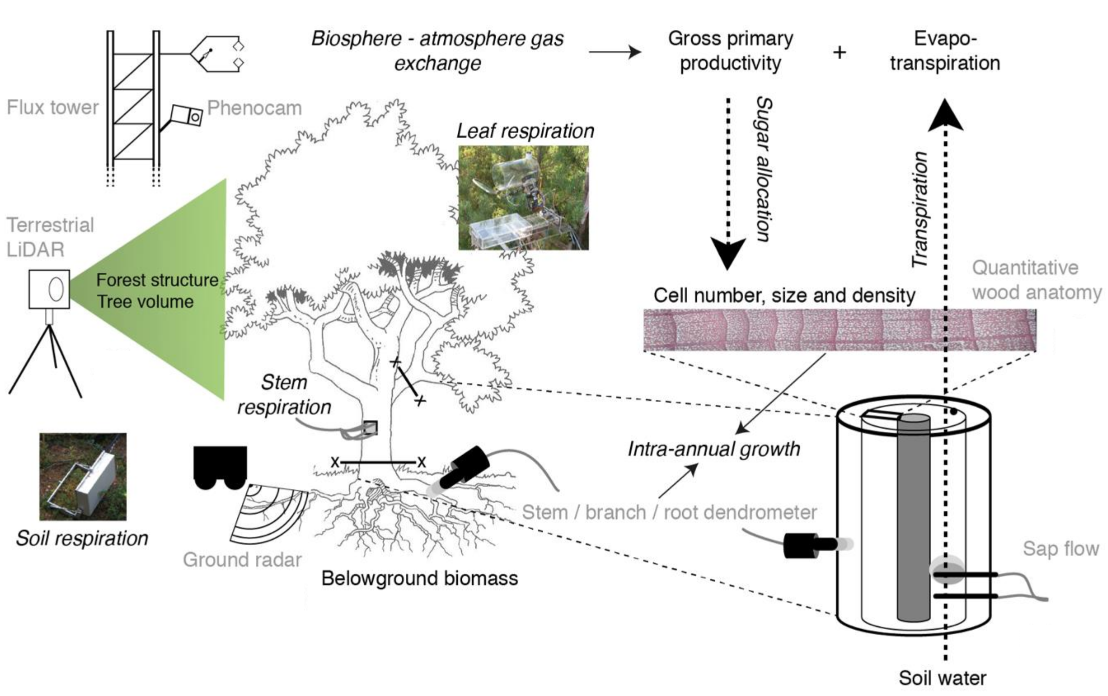
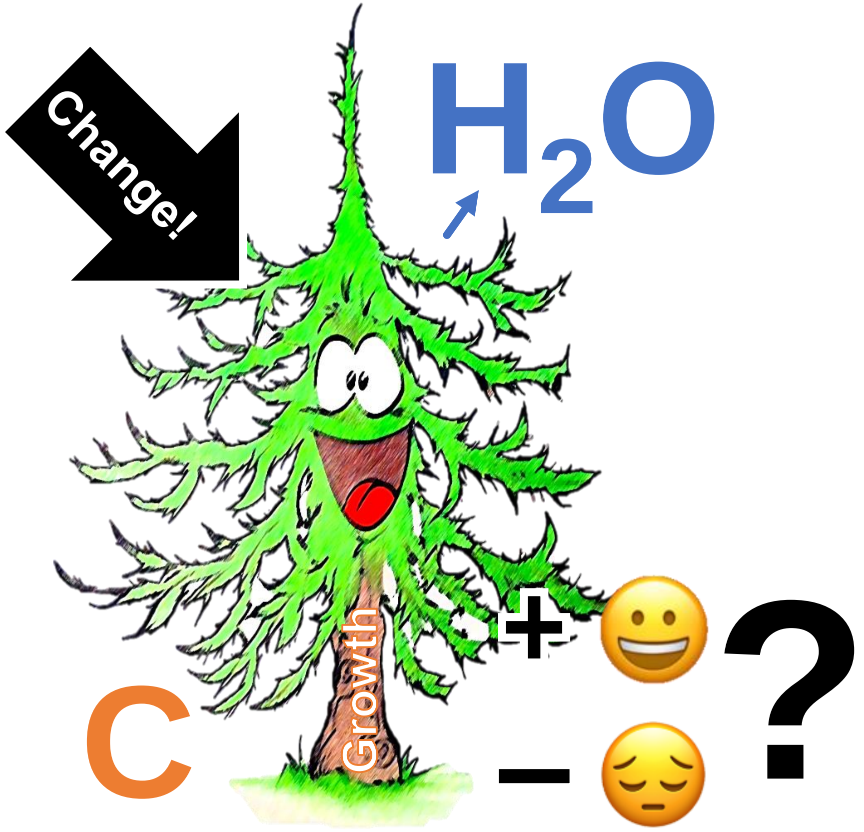


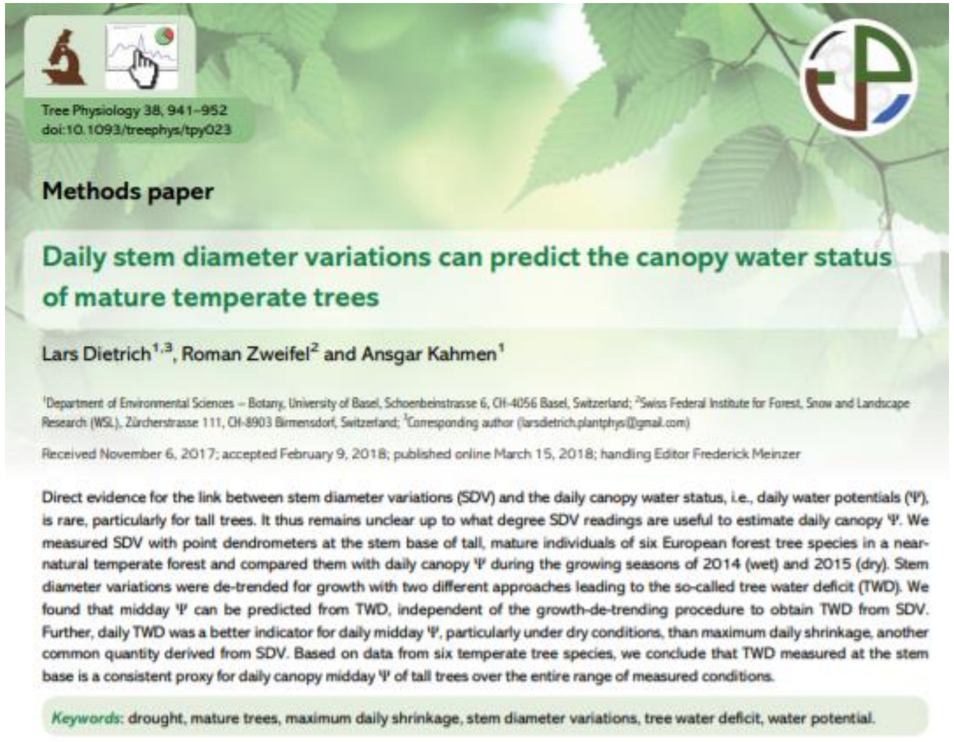
Fig. Schematic of monitoring sites which aim at consistent observations of processes that determine carbon allocation and water use.

Source: Babst *et al.* (2021) doi: https://doi.org/10.1016/j.tplants.2020.10.002



Monitoring examples

Dendrometers *Stem radius changes of trees provide information on canopy water status*



Tree Physiology 38, 941–952
doi:10.1093/treephys/tpy023

Methods paper

Daily stem diameter variations can predict the canopy water status of mature temperate trees

Lars Dietrich[1,3], Roman Zweifel[2] and Ansgar Kahmen[1]

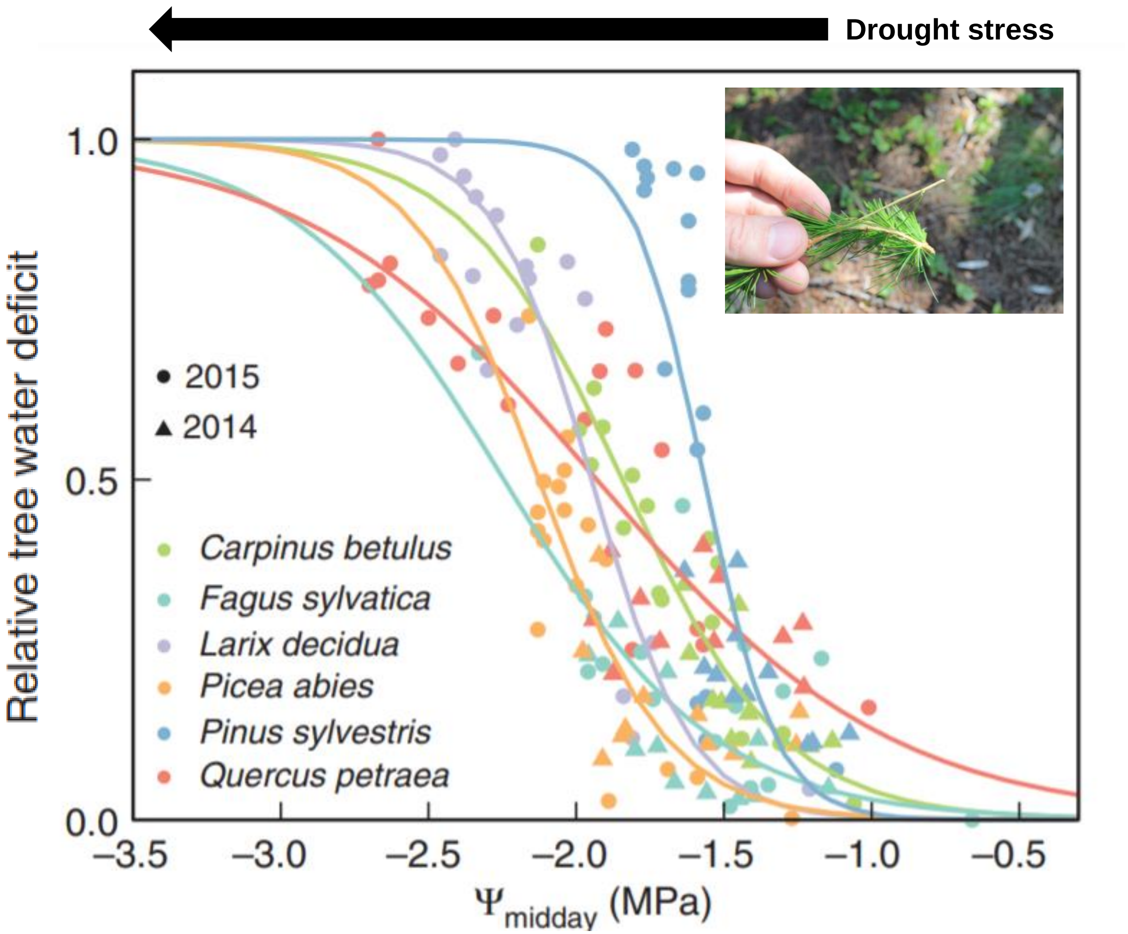
[1]Department of Environmental Sciences – Botany, University of Basel, Schoenbeinstrasse 6, CH-4056 Basel, Switzerland; [2]Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland; [3]Corresponding author (larsdietrich.plantphys@gmail.com)

Received November 6, 2017; accepted February 9, 2018; published online March 15, 2018; handling Editor Frederick Meinzer

Direct evidence for the link between stem diameter variations (SDV) and the daily canopy water status, i.e., daily water potentials (Ψ), is rare, particularly for tall trees. It thus remains unclear up to what degree SDV readings are useful to estimate daily canopy Ψ. We measured SDV with point dendrometers at the stem base of tall, mature individuals of six European forest tree species in a near-natural temperate forest and compared them with daily canopy Ψ during the growing seasons of 2014 (wet) and 2015 (dry). Stem diameter variations were de-trended for growth with two different approaches leading to the so-called tree water deficit (TWD). We found that midday Ψ can be predicted from TWD, independent of the growth-de-trending procedure to obtain TWD from SDV. Further, daily TWD was a better indicator for daily midday Ψ, particularly under dry conditions, than maximum daily shrinkage, another common quantity derived from SDV. Based on data from six temperate tree species, we conclude that TWD measured at the stem base is a consistent proxy for daily canopy midday Ψ of tall trees over the entire range of measured conditions.

Keywords: drought, mature trees, maximum daily shrinkage, stem diameter variations, tree water deficit, water potential.

Fig. Example of relevant literature related to using dendrometer measurements.



Monitoring examples

Dendrometers *Structure of the treenetproc R package and accessibility*

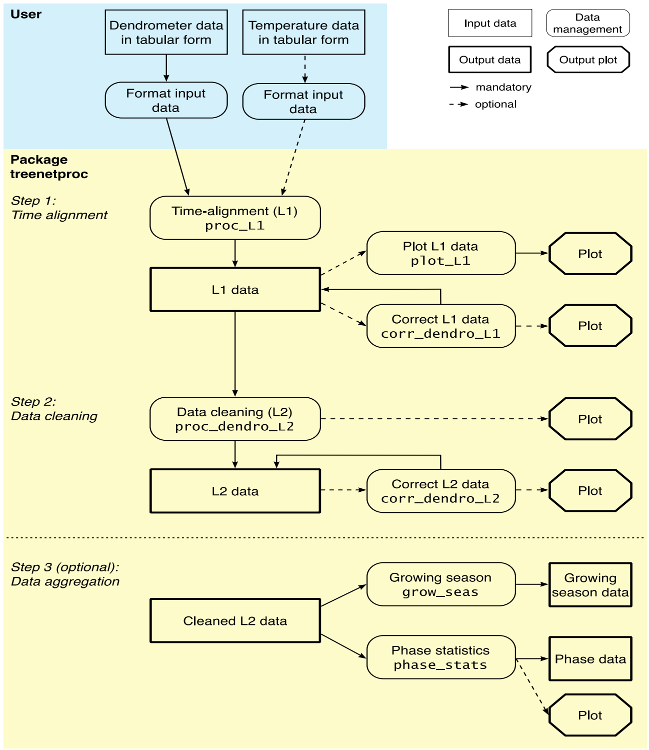


Fig. Workflow of treenetproc. Functions of the package are shown in rounded rectangles, including a description of the functionality (first line) and the name of the function (second line). Outputs of the functions are shown as rectangles (data output) and octagons (graphical output). Solid arrows indicate the regular workflow, dashed arrows indicate the optional, user-induced workflow.

Availability

The R package treenetproc and all source code is available on GitHub (https://github.com/treenet/treenetproc). In the R software (Team, 2019), the package can be installed with the following commands:

```
# install.packages("devtools")
library(devtools)
devtools::install_github("treenet/treenetproc")
```

Link

An R tutorial on the use of treenetproc is accessible via this link:
https://deep-tools.netlify.app/2020/11/21/treenetproc-intro/

Structure

The general workflow of treenetproc is composed of three main steps including multiple functions [Fig.]. In step 1 (time alignment), the raw data of dendrometer is aligned to user-defined, regular time steps (L1 data). In step 2 (data cleaning), outliers and shifts in the L1 data are detected and corrected. In step 3, the L2 data is analyzed, and several derived variables are calculated.

References

Zweifel *et al.* (2016) doi: https://doi.org/10.1111/nph.13995

Monitoring examples

Sap flow *Thermal dissipation probes can extract stomatal behavior*


Tree Physiology 38, 953–964
doi:10.1093/treephys/tpy043

Research paper

Boreal tree hydrodynamics: asynchronous, diverging, yet complementary

Christoforos Pappas[1,10], Ashley M. Matheny[2,3], Jennifer L. Baltzer[4], Alan G. Barr[5], T. Andrew Black[6], Gil Bohrer[3], Matteo Detto[7,8], Jason Maillet[9], Alexandre Roy[1], Oliver Sonnentag[1] and Jilmarie Stephens[6]

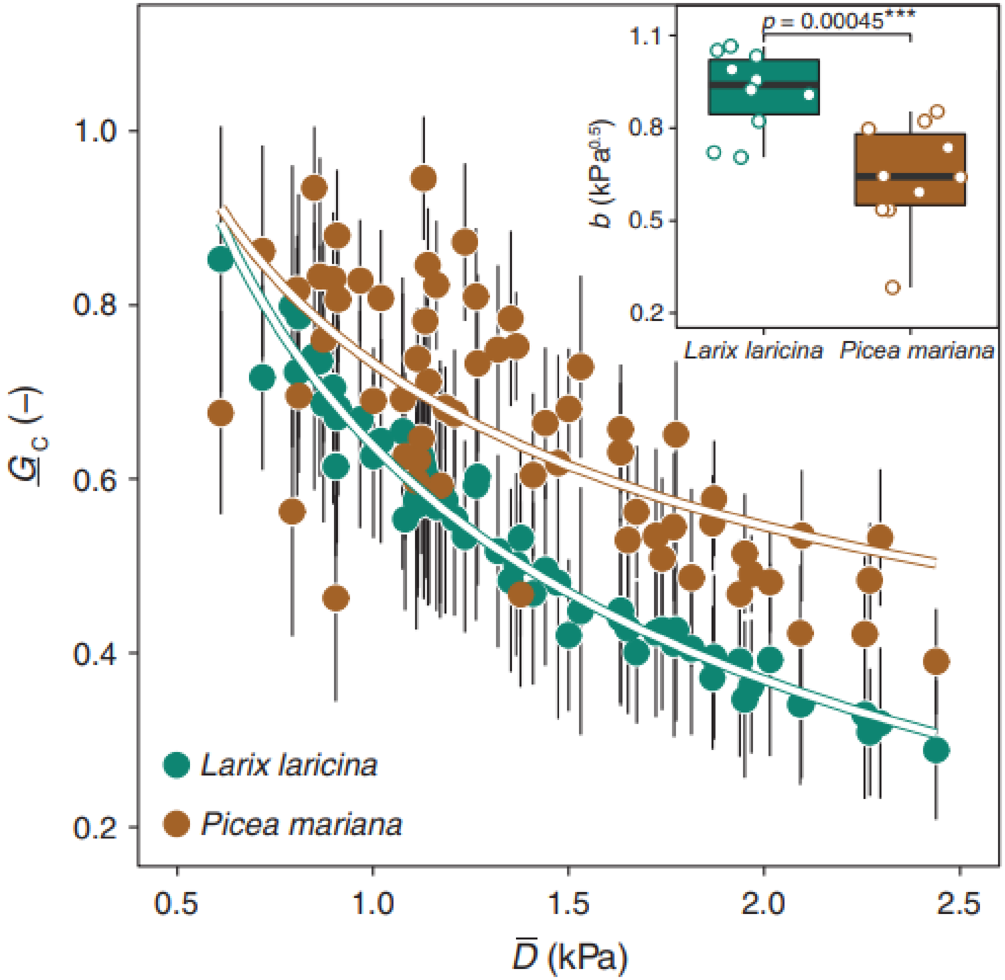
[1]Département de géographie and Centre d'études nordiques, Université de Montréal, Montréal, QC H2V 2B8, Canada; [2]Department of Geological Sciences, Jackson School of Geosciences, University of Texas at Austin, Austin, TX 78712, USA; [3]Department of Civil, Environmental, and Geodetic Engineering, Ohio State University, Columbus, OH 43210, USA; [4]Department of Biology, Wilfrid Laurier University, Waterloo, ON N2L 3C5, Canada; [5]Climate Research Division, Environment Canada and Global Institute for Water Security, University of Saskatchewan, Saskatoon, SK S7N 3H5, Canada; [6]Faculty of Land and Food Systems, Biometeorology and Soil Physics Group, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; [7]Center for Tropical Forest Science-Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Panamá 0843-03092, Republic of Panamá; [8]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA; [9]School of Environment and Sustainability, University of Saskatchewan, Saskatoon, SK S7N 5C8, Canada; [10]Corresponding author (christoforos.pappas@umontreal.ca)

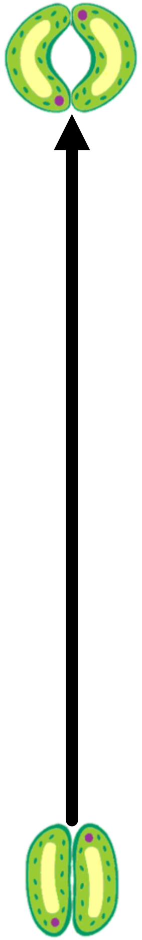
Received October 18, 2017; accepted April 17, 2018; published online May 8, 2018; handling Editor Nathan Phillips

Water stress has been identified as a key mechanism of the contemporary increase in tree mortality rates in northwestern North America. However, a detailed analysis of boreal tree hydrodynamics and their interspecific differences is still lacking. Here we examine the hydraulic behaviour of co-occurring larch (*Larix laricina*) and black spruce (*Picea mariana*), two characteristic boreal tree species, near the southern limit of the boreal ecozone in central Canada. Sap flux density (J_s), concurrently recorded stem radius fluctuations and meteorological conditions are used to quantify tree hydraulic functioning and to scrutinize tree water-use strategies. Our analysis revealed asynchrony in the diel hydrodynamics of the two species with the initial rise in J_s occurring 2 h earlier in larch than in black spruce. Interspecific differences in larch and black spruce crown architecture explained the observed asynchrony in their hydraulic functioning. Furthermore, the two species exhibited diverging stomatal regulation strategies with larch and black spruce employing relatively isohydric and anisohydric behaviour, respectively. Such asynchronous and diverging tree-level hydrodynamics provide new insights into the ecosystem-level complementarity in tree form and function, with implications for understanding boreal forests' water and carbon dynamics and their resilience to environmental stress.

Keywords: boreal forest, crown complementarity, dendrometers, isohydricity, *Larix laricina* (larch, tamarack), *Picea mariana* (black spruce), plant hydraulics, sap flow.

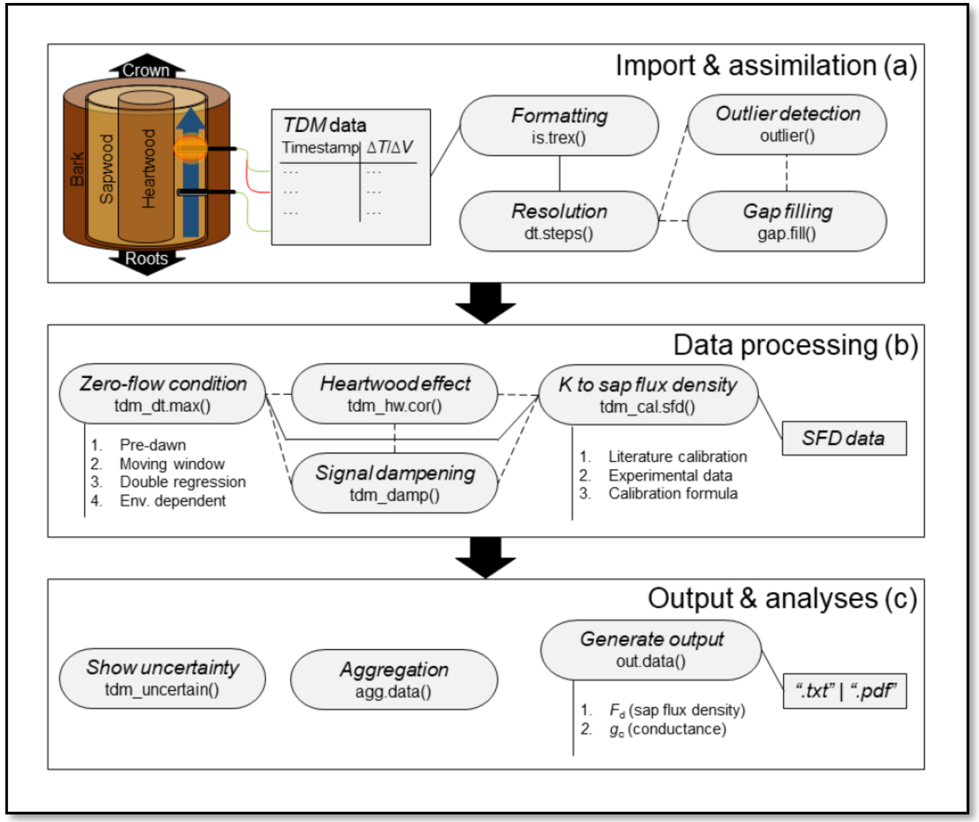
Fig. Example of relevant literature related to using sap flow measurements.





Monitoring examples

Sap flow *Structure of the TREX R package and accessibility*



Availability

The R package TREX and all source code is available on GitHub (https://the-hull.github.io/TREX). In the R software (Team, 2019), the package can be installed with the following commands:

```
# install.packages("devtools")
library(devtools)
devtools::install_github("the-Hull/TREX")
```

Link

An R tutorial on the use of treenetproc is accessible via this link: https://deep-tools.netlify.app/2020/11/23/trex-intro/

Structure

The general workflow of TREX is composed of three main steps including multiple functions [Fig.]. In step 1 (import & assimilation), the raw sap flow data and the associated auxiliary meteorological data are imported in R and the consistency of the time series object is tested and if necessary corrected (i.e., regular time steps of time series objects, outlier detection, gap filling). In step 2 (data processing), zero-flow conditions can be derived with several approaches, and corrections can be applied (i.e., to heartwood correction and dampening). Then, sap flux density can be estimated using user-specific or literature values of the calibration parameters. In step 3 (output & analyses), the uncertainties associated with the sap flow pre-processing assumptions can be quantified with state-of-the-art statistical methods, the temporal resolution of the generated data can be adjusted, and the crown conductance to water can be estimated

References

Peters *et al.* 2020 doi: https://doi.org/10.1111/2041-210X.13524