

# Tree growth rates: Data preparation

John L. Godlee



THE UNIVERSITY of EDINBURGH  
School of GeoSciences



# Issues installing packages

---

Some people had trouble installing `ProdVital` and `seosawr`

Sharepoint:

```
rscripts/seosawr_fn.R
```

```
rscripts/ProdVital_fn.R
```

Instead of running `library(seosawr)` and `library(ProdVital)`:

```
source("rscripts/seosawr_fn.R")
```

```
source("rscripts/ProdVital_fn.R")
```

# Data preparation

1. Remove measurements of dead stems
2. Identify outlier measurements
3. Account for changes in POM
4. Standardise species names

Sharepoint: "worksheets/2\_cleaning\_2023-09-01.html"

## SEOSAW Carbon Dynamics Workshop: Part 2

Data preparation: mortality, outliers, species names

John L. Godlee

2023-08-25

- [Introduction](#)
- [Stem mortality timelines](#)
- [Outlier measurements](#)
- [POM](#)
- [Duplicate measurements](#)
- [Species names](#)
- [Finishing off](#)

### Introduction

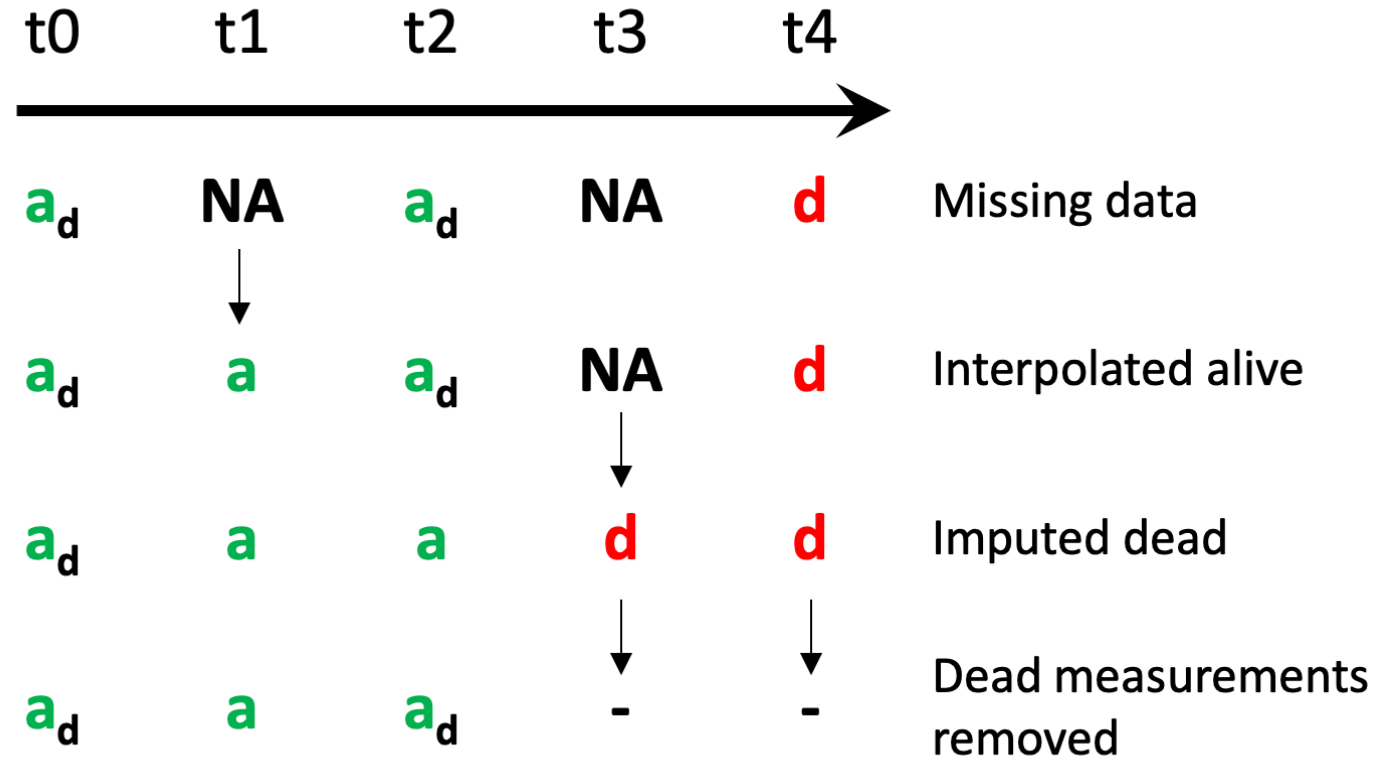
This worksheet provides examples of how to prepare repeat stem measurement data to measure carbon dynamics. The code chunks are all written in R. The example dataset used in this worksheet ("stems\_ABG.csv") can be found in the [workshop Sharepoint folder](#).

First, load the required packages. If you haven't installed the packages you should do so with `install.packages("<pkg>")`, where `<pkg>` is the name of the package.

# Mortality timelines

```
# Check and impute mortality
mor <- statusImputGen(s)
s_mor <- bind_cols(s, mor)

# Remove dead measurements
s_a <- s_mor %>%
  filter(stem_status_est == "a")
```

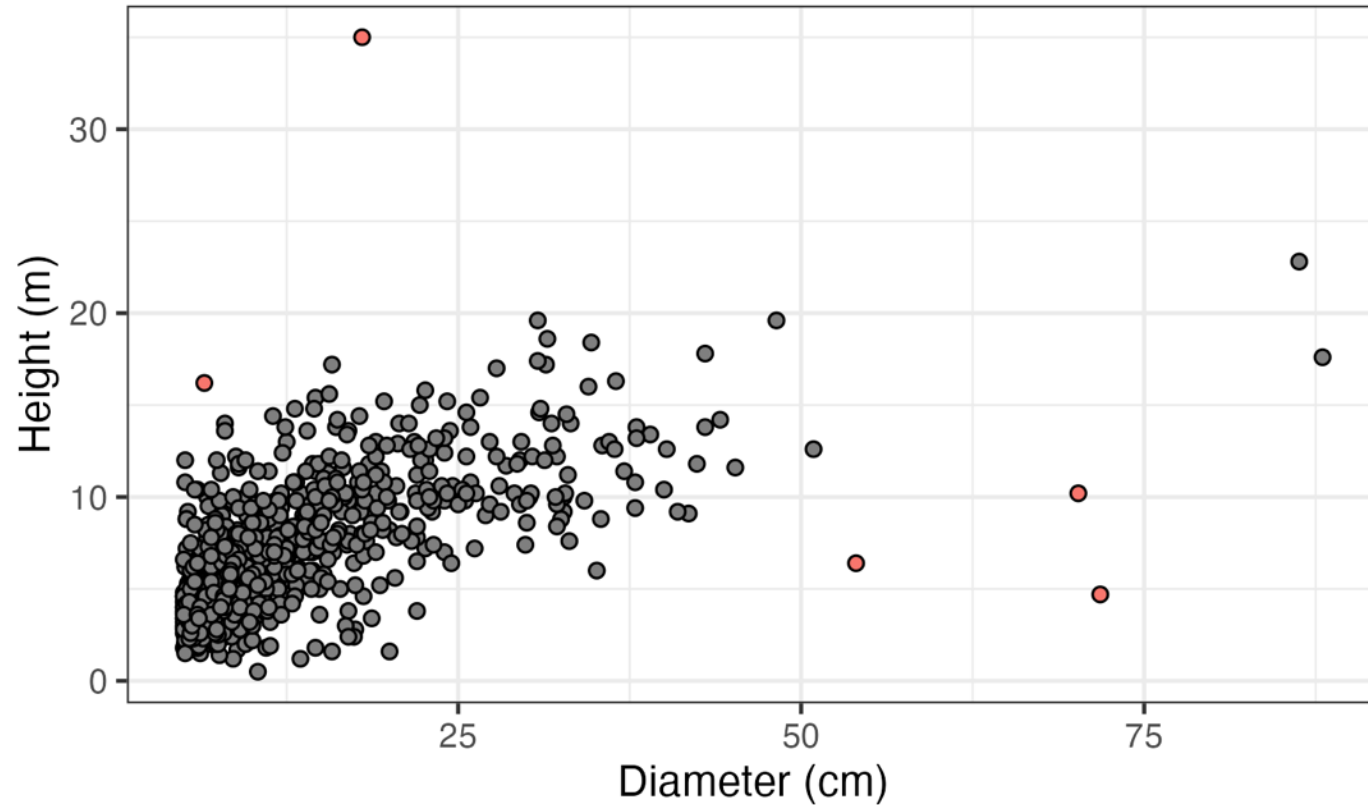


# Identify outliers

```
# Identify outlier diameter measurements  
dflag <- diamHeightCheck(s_a)[[1]]
```

```
# Join outliers to data  
dflag$flag <- TRUE  
s_d <- s %>%  
  left_join(., dflag) %>%  
  mutate(dflag = ifelse(  
    is.na(dflag), FALSE, dflag))
```

```
# Plot data, highlight outliers  
ggplot(s_d,  
  aes(x = diam, y = height)) +  
  geom_point(aes(colour = flag))
```

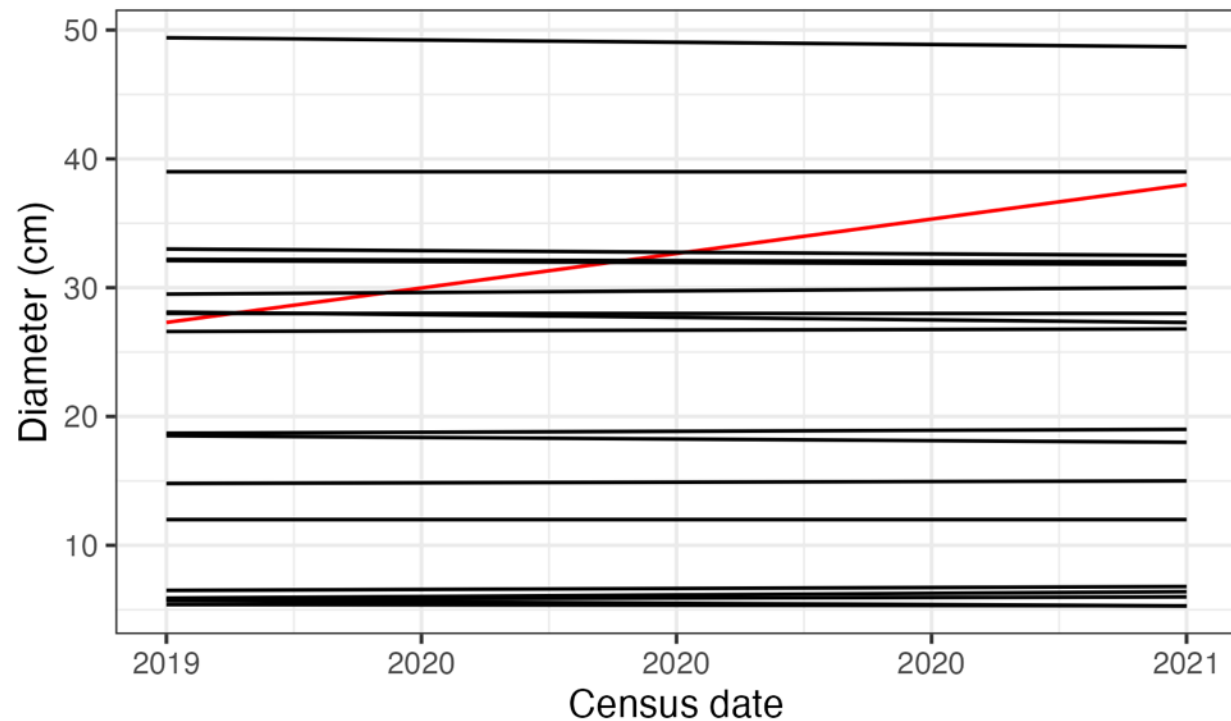


# Identify outliers in timelines

```
# Identify outlier diameter measurements  
dcflag <- diamChangeCheck(s, thresh = 4)[[1]]
```

```
# Join outliers to data  
dcflag$flag <- TRUE  
s_dc <- s %>%  
  left_join(., dcflag)  
  mutate(dflag = ifelse(  
    is.na(dflag), FALSE, dflag))
```

```
# Plot data, highlight outliers  
ggplot(s_dc,  
  aes(x = census_date, y = diam)) +  
  geom_line(aes(colour = flag,  
    group = stem_id))
```



# Account for POM changes

---

Trees taper upwards.

If POM is moved, diameter increment will be biased.

Either:

- exclude stems where the POM changes
- estimate diameter at original POM

```
# Fill missing POM values
pom_imp <- pomInputGen(s_dc, p)
s_dc$pom_est <- pom_imp$pom_est

# Adjust diameter where POM has moved
s$diam_adj <- diamAdj(s, pom = "pom_est")

# Exclude stems where POM has moved
s_pom <- s %>%
  group_by(plot_id, stem_id) %>%
  filter(n_distinct(pom_est) == 1)
```

# Check measurements are unique

---

Each stem should only be measured once per census

```
# Flag duplicate measurements  
seosawr::measurementUniqueCheck(s_pom)
```

Returns a dataframe of IDs where a stem is measured more than once per census.

# Standardise species names

---

Species used to estimate wood density, for biomass calculations.

Consistent names needed if subsetting growth rates by species.

Assume species names are refined over consecutive censuses.

Use `species_name_clean`  
(SEOSAW standardised taxonomy)

```
# Backfill species names
s_pom$species_fill <-
  speciesFill(s,
    species = "species_name_clean")
```

# Data preparation

1. Remove measurements of dead stems
2. Identify outlier measurements
3. Account for changes in POM
4. Standardise species names

Sharepoint: "worksheets/2\_cleaning\_2023-09-01.html"

## SEOSAW Carbon Dynamics Workshop: Part 2

Data preparation: mortality, outliers, species names

John L. Godlee

2023-08-25

- [Introduction](#)
- [Stem mortality timelines](#)
- [Outlier measurements](#)
- [POM](#)
- [Duplicate measurements](#)
- [Species names](#)
- [Finishing off](#)

### Introduction

This worksheet provides examples of how to prepare repeat stem measurement data to measure carbon dynamics. The code chunks are all written in R. The example dataset used in this worksheet ("stems\_ABG.csv") can be found in the [workshop Sharepoint folder](#).

First, load the required packages. If you haven't installed the packages you should do so with `install.packages("<pkg>")`, where `<pkg>` is the name of the package.