universität freiburg

05 – Review of Inventory and Uncertainties

Geosense
Maximilian Fabi

Agenda What's happening today

1. Review of last week's inventory
2. Introduction to sampling uncertainty
3. Transition to the exercise in R

Results of our Inventory → Sampling uncertainty → Simulation of sampling a Forest

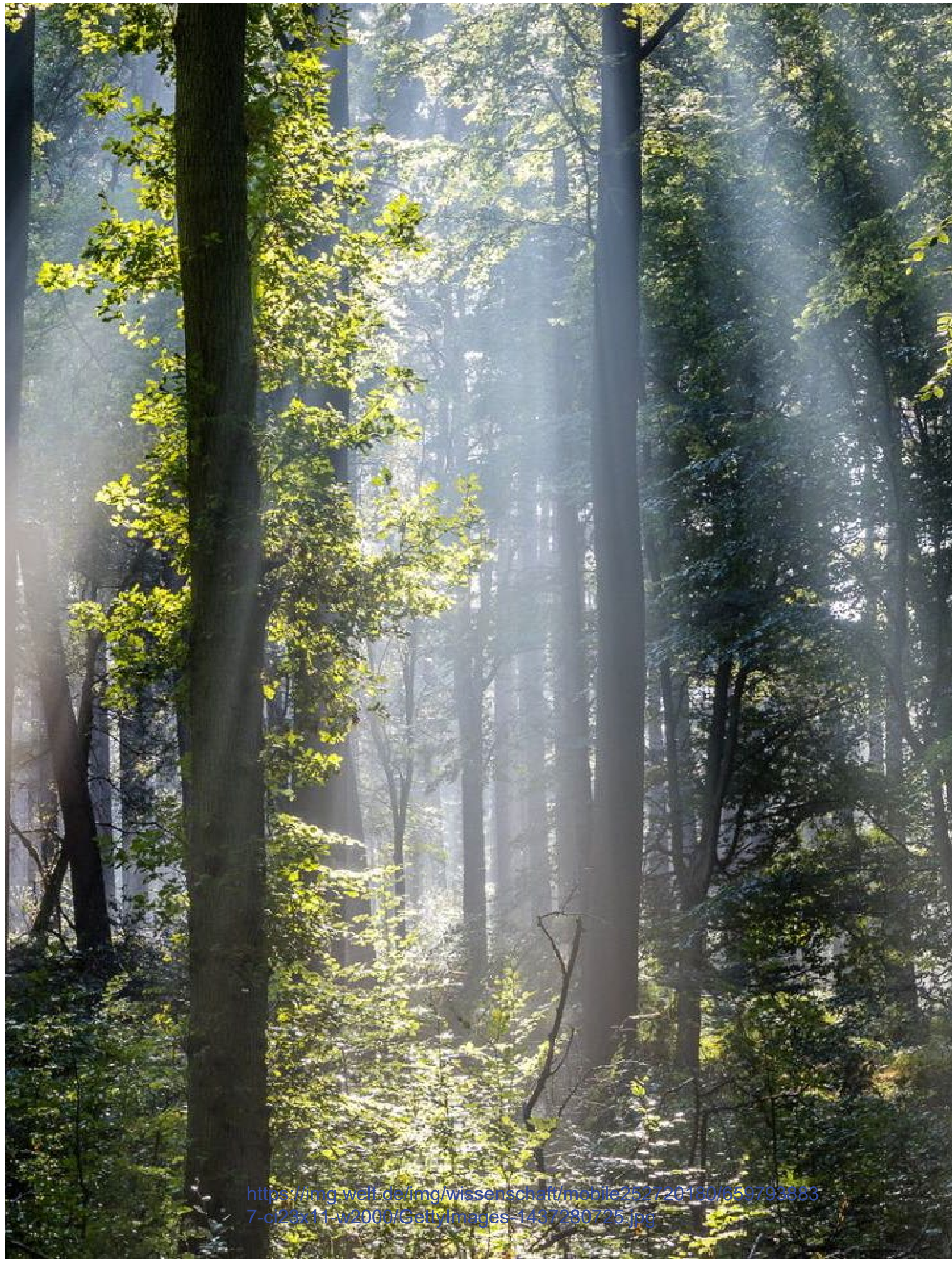
review

The data you recorded and handed in...

Objectives of the field exercise

- Getting to know the techniques and uncertainties
- Influence of the correct selection of plots
- Human factor
- Measurement inaccuracy of the instruments used
- Influence of the weather ...

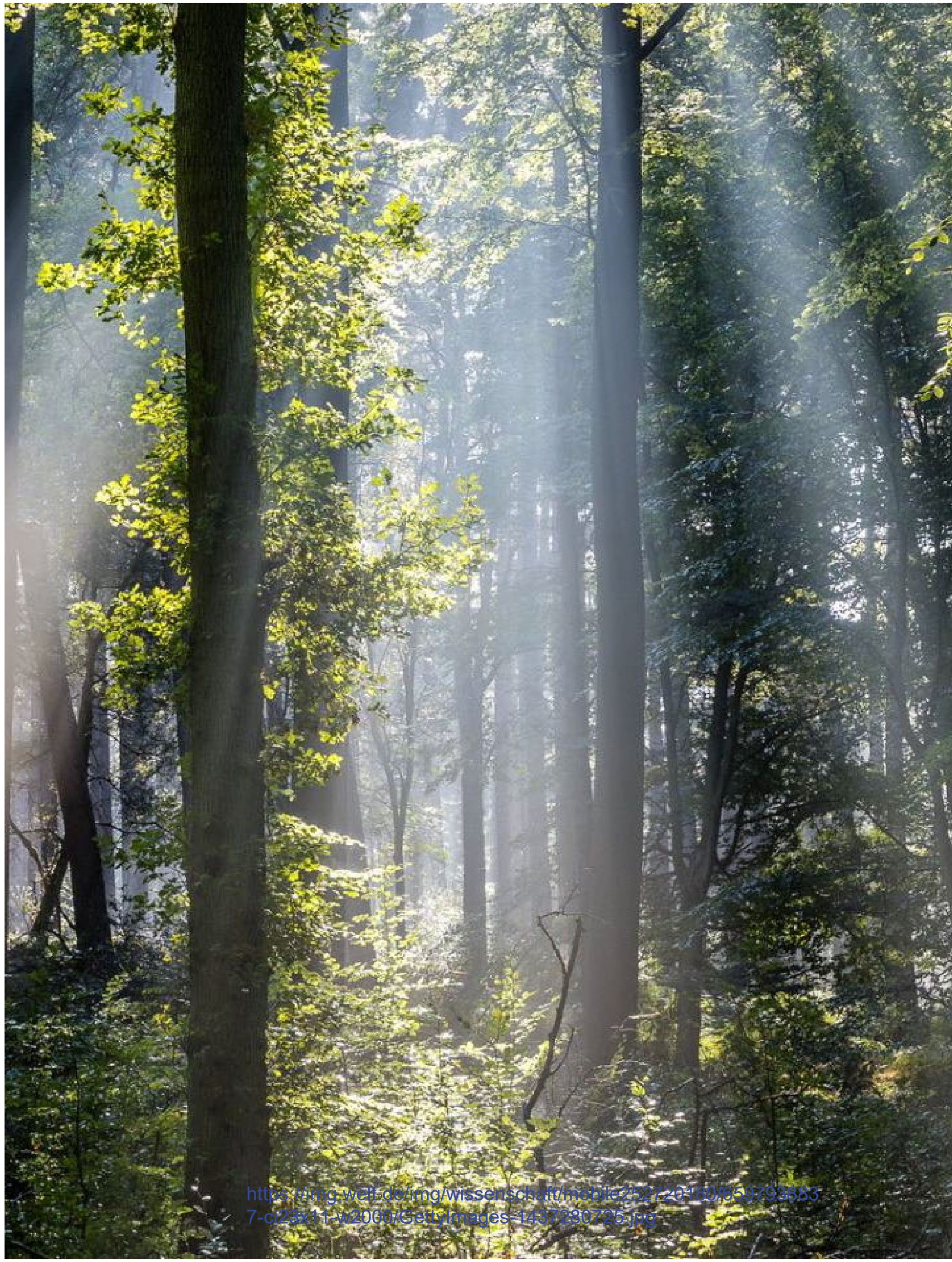
- AND: Learning to submit data according to specifications is also included here! Sadly only one group managed to do that properly
  - Anglecount missing, plot ID missing, coordinates missing…



https://img.welt.de/img/wissenschaft/mobile252720180/6597938837-ci23x11-w2000/GettyImages-1437280725.jpg

Results

- What could we conclude from the collected data?

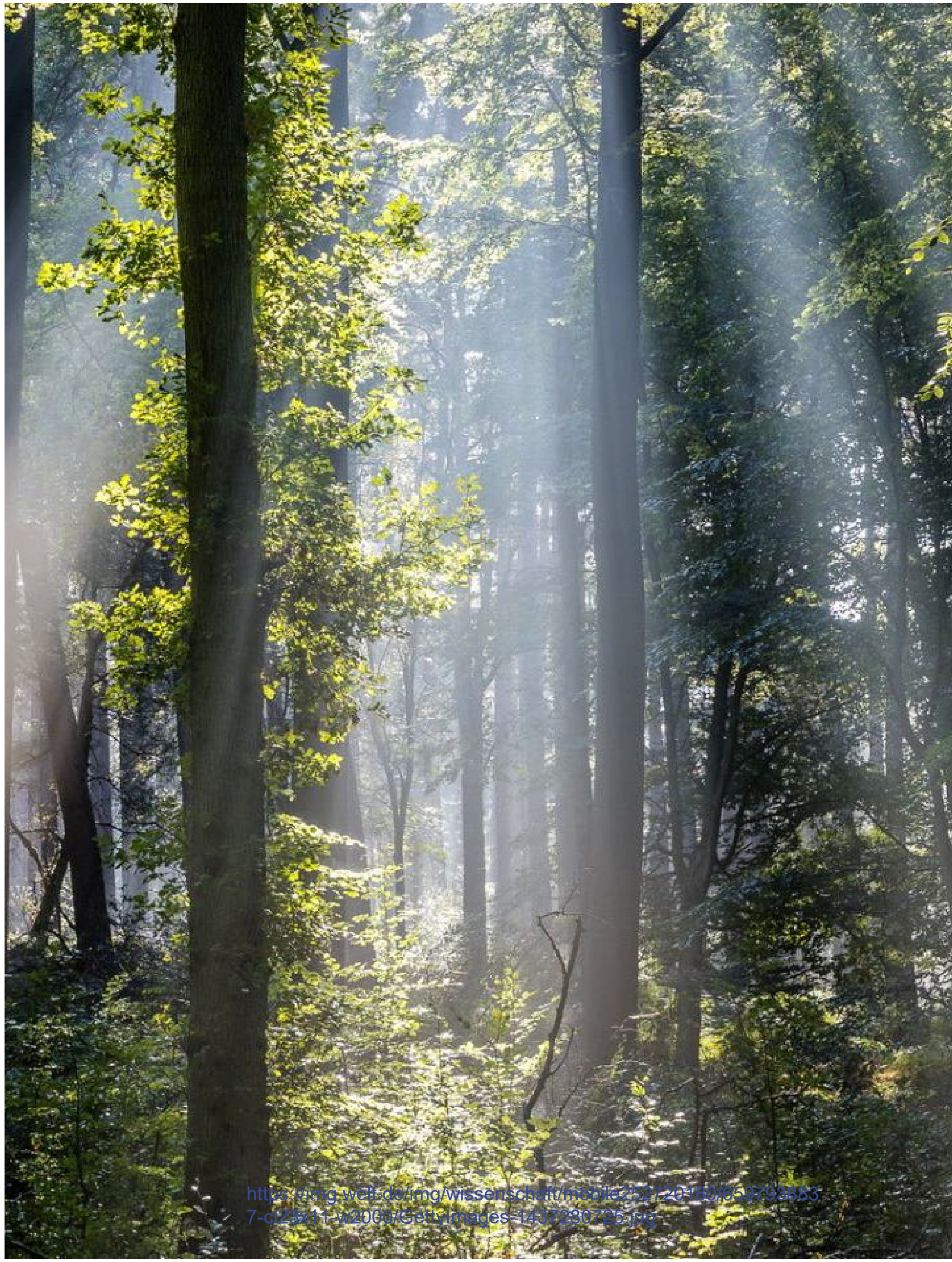


https://img.welt.de/img/wissenschaft/mobile252720180/6597938837-ci23x11-w2000/GettyImages-1437280725.jpg

Results

- Basal area and volume

- Basal area: Sum of the basal areas of all trees on a specific area (m2/ha)
- Volume: Cubic meters per hectar (m3/ha)
- Proportions of tree species, biomass, …



https://img.welt.de/img/wissenschaft/mobile252720180/6597938837-ci23x11-w2000/GettyImages-1437280725.jpg

Results

Mooswald Inventory 2024

Plot Nr:7

Regeneration plots 0-2m radius

		Q1 (N > E)			Q2 (E > S)			Q3 (S > W)			Q4 (W > N)		
	Max Number of trees to be counted per height class	1	4	25	1	4	25	1	4	25	1	4	25
		61A	61B	61C	62A	62B	62C	63A	63B	63C	64A	64B	64C
	Number of trees (all species)	0	3	2	0	2	1	0	0	1	0	0	0
Pos.	**Species** [70]	71A	71B	71C	72A	72B	72C	73A	73B	73C	74A	74B	74C
1	Eiche		33%										
2	Roteiche			100%		50%				100%			
3	Hainbuche		66%										
4	Spitzahorn					50%	100%						
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													

Single tree measurements from DBH > 7.0 cm

Pos	Tree ID [40]	Species [41]	Azimuth [42]	Distance [43]	DBH [44]	DBH height [45	Height [46]
1	1	Hainbuche	230°	2,5m	14cm		10,5m
2	2	Bergahorn	220°	4,4m	25cm		19m
3	3	Bergahorn	340°	2,8m	33cm		21m
4	4	Linde	170°	5,5m	35,3cm		24m
5	5	Bergahorn	130°	8,3m	38cm		25,4m
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							

Angle count survey

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Average
m2/ha	28	28				

What do we need for the basal area and the volume?

Plot Nr: 8

Regeneration

		Q1 (N > E)			Q2 (E > S)			Q3 (S > W)			Q4 (W > N)		
	Max Number of trees to be counted per height class	1	4	25	1	4	25	1	4	25	1	4	25
		61A	61B	61C	62A	62B	62C	63A	63B	63C	64A	64B	64C
	Number of trees (all species)	1	0	0	1	0	0	1	0	0	0	0	0
Pos.	**Species** [70]	71A	71B	71C	72A	72B	72C	73A	73B	73C	74A	74B	74C
1	Roteiche	50%											
2	Hainbuche	50%			100%			100%					
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													

Single tree measurements from DBH > 7.0 cm

Pos	Tree ID [40]	Species [41]	Azimuth [42]	Distance [43]	DBH [44]	DBH height [45	Height [46]
1	1	Stileiche	355°	1,5m	25,3cm		18,5m
2	2	Bergahorn	260°	2,3m	19cm		17m
3	3	Linde	45°	4,9m	15cm		6m
4							
5	4	Spitzahorn	170°	3,5m	33cm		25m
6	5	Stileiche	270°	6,2m	38cm		20m
7	6	Hainbuche	320°	6,4m	31cm		17,5m
8	7	Bergahorn	60°	10,2m	31cm		26m
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							

Angle count survey

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Average
m2/ha	28	28				

Plot Nr: 1

Regeneration plots 0-2m radius

		Q1 (N > E)			Q2 (E > S)			Q3 (S > W)			Q4 (W > N)		
	Max Number of trees to be counted per height class	1	4	25	1	4	25	1	4	25	1	4	25
		61A	61B	61C	62A	62B	62C	63A	63B	63C	64A	64B	64C
	Number of trees (all species)	0	2	5	0	0	12	0	0	17	0	2	18
Pos.	**Species** [70]	71A	71B	71C	72A	72B	72C	73A	73B	73C	74A	74B	74C
1	Roteiche		100%	100%			100%			100%		100%	100%
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													

Single tree measurements from DBH > 7.0 cm

Pos	Tree ID [40]	Species [41]	Azimuth [42]	Distance [43]	DBH [44]	DBH height [45	Height [46]
1	1	Hainbuche	155°	2,8m	10cm		7m
2	2	(Winter-)Linde	45°	3m	12,5cm		17m
3							
4	3	Hainbuche	240°	3,3m	21cm		18m
5	4	(Winter-)Linde	160°	3,8m	23cm		20m
6							
7	5	Bergahorn	80°	2,9m	34cm		24m
8	6	(Sommer-)Linde	30°	10,6m	34cm		24m
9	7	Kirsche	330°	11,9m	41cm		29m
10	8	Stieleiche	310°	7,8m	32,5cm		24m
11	9	Linde	300°	3m	40cm		26m
12	10	Roteiche	270°	10,2m	60,5cm		35m
13	11	Linde	180°	8,3m	44cm		28m
14							
15							
16							
17							
18							
19							
20							

Angle count survey

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Average
m2/ha	34					

Results

Mooswald Inventory 2024

Plot Nr:7 — Regeneration plots 0-2m radius

		Q1 (N > E)			Q2 (E > S)			Q3 (S > W)			Q4 (W > N)		
	Max Number of trees to be counted per height class	1	4	25	1	4	25	1	4	25	1	4	25
		61A	61B	61C	62A	62B	62C	63A	63B	63C	64A	64B	64C
	Number of trees (all species)	0	3	2	0	2	1	0	0	1	0	0	0
Pos.	Species [70]	71A	71B	71C	72A	72B	72C	73A	73B	73C	74A	74B	74C
1	Eiche		33%										
2	Roteiche			100%		50%				100%			
3	Hainbuche		66%										
4	Spitzahorn					50%	100%						
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													

Single tree measurements from DBH > 7.0 cm

Pos	Tree ID [40]	Species [41]	Azimuth [42]	Distance [43]	DBH [44]	DBH height [45]	Height [46]
1	1	Hainbuche	230°	2,5m	14cm		10,5m
2	2	Bergahorn	220°	4,4m	25cm		19m
3	3	Bergahorn	340°	2,8m	33cm		21m
4	4	Linde	170°	5,5m	35,3cm		24m
5	5	Bergahorn	130°	8,3m	38cm		25,4m
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							

Angle count survey

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Average
m2/ha	28	28				

Plot Nr: 8 — Regenerat…

													25
	Max Number of… to be counted… height clas…												64C
	Number of trees… species)											0	0
Pos.	Species												74C
1	Roteiche												
2	Hainbuche												
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													

Single tree measurements from DBH > 7.0 cm

Pos	Tree ID [40]	Species [41]	Azimuth [42]	Distance [43]	DBH [44]	DBH height [45]	Height [46]
1	1	Stileiche	355°	1,5m	25,3cm		18,5m
2	2	Bergahorn	260°	2,3m	19cm		17m
3	3	Linde	45°	4,9m	15cm		6m
4							
5	4	Spitzahorn	170°	3,5m	33cm		25m
6	5	Stileiche	270°	6,2m	38cm		20m
7	6	Hainbuche	320°	6,4m	31cm		17,5m
8	7	Bergahorn	60°	10,2m	31cm		26m
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							

Angle count survey

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Average
m2/ha	28	28				

Plot Nr: 1 — Regeneration plots 0-2m radius

		Q1 (N > E)			Q2 (E > S)			Q3 (S > W)			Q4 (W > N)		
	Max Number of trees to be counted per height class	1	4	25	1	4	25	1	4	25	1	4	25
		61A	61B	61C	62A	62B	62C	63A	63B	63C	64A	64B	64C
	Number of trees (all species)	0	2	5	0	0	12	0	0	17	0	2	18
Pos.	Species [70]	71A	71B	71C	72A	72B	72C	73A	73B	73C	74A	74B	74C
1	Roteiche		100%	100%			100%			100%		100%	100%
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													

Single tree measurements from DBH > 7.0 cm

Pos	Tree ID [40]	Species [41]	Azimuth [42]	Distance [43]	DBH [44]	DBH height [45]	Height [46]
1	1	Hainbuche	155°	2,8m	10cm		7m
2	2	(Winter-)Linde	45°	3m	12,5cm		17m
3							
4	3	Hainbuche	240°	3,3m	21cm		18m
5	4	(Winter-)Linde	160°	3,8m	23cm		20m
6							
7	5	Bergahorn	80°	2,9m	34cm		24m
8	6	(Sommer-)Linde	30°	10,6m	34cm		24m
9	7	Kirsche	330°	11,9m	41cm		29m
10	8	Stieleiche	310°	7,8m	32,5cm		24m
11	9	Linde	300°	3m	40cm		26m
12	10	Roteiche	270°	10,2m	60,5cm		35m
13	11	Linde	180°	8,3m	44cm		28m
14							
15							
16							
17							
18							
19							
20							

Angle count survey

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Average
m2/ha	34					

- Basal area:

$$\left(\frac{\frac{DBH}{100}}{2}\right)^2 * \pi$$

- Volume:

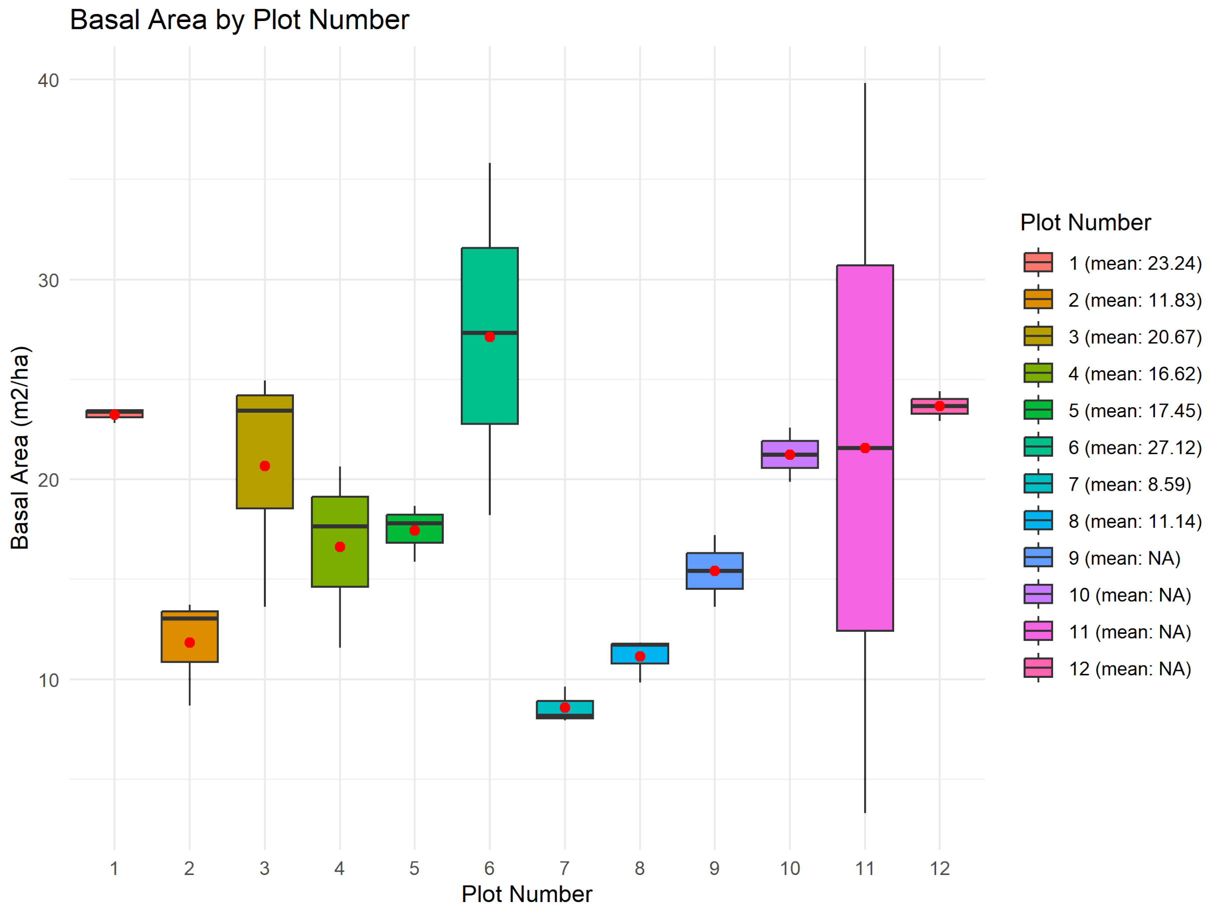
$$height * basalarea * 0.5$$

Results

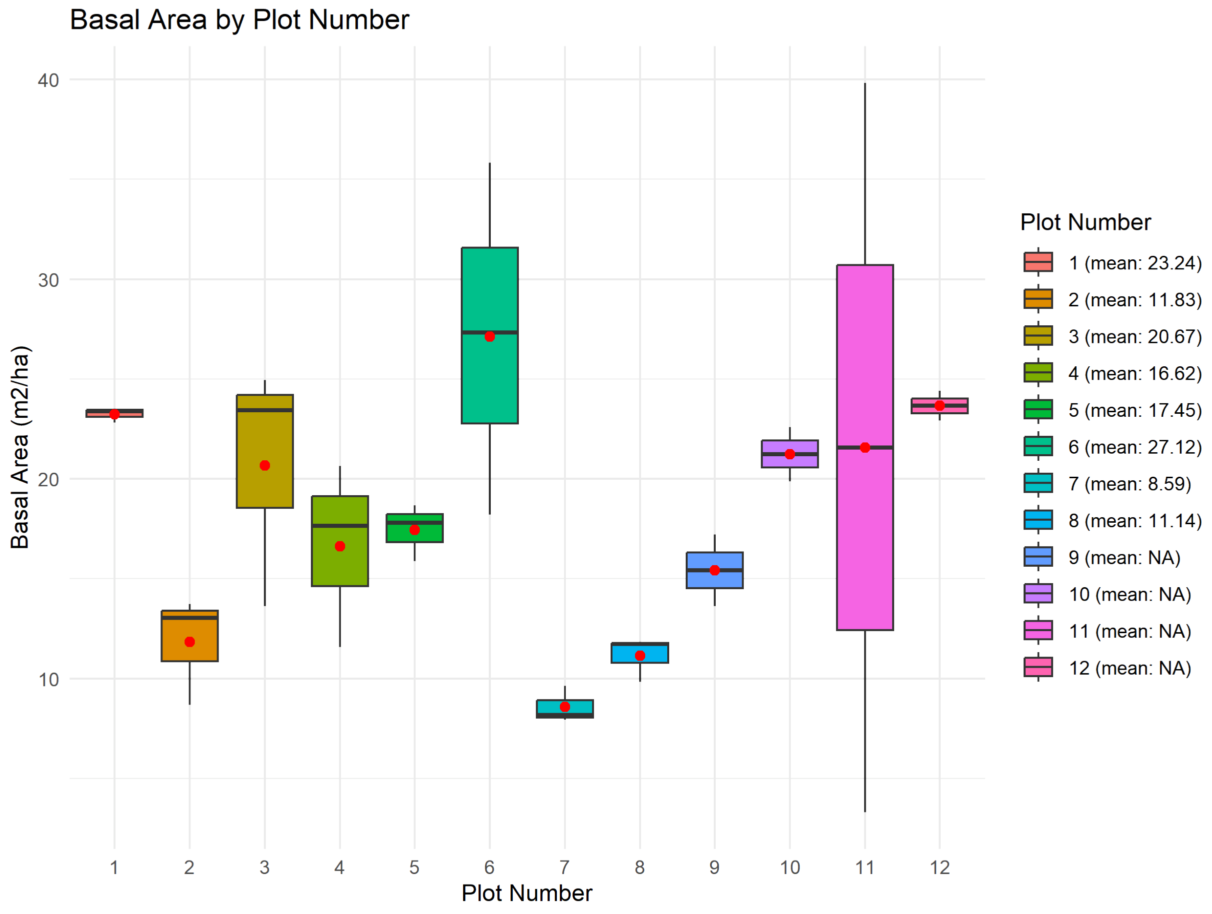
plot_id	basal_area[m2/ha]	volume[m3]	angle_count[m2/ha]
1_1	22.81	15.79	29
1_2	23.39	14.74	34
1_3	23.51	13.75	24
2_1	13.05	7.08	31
2_2	8.69	3.91	
2_3	13.74	7.24	32
3_1	23.44	14.53	28
3_2	24.94	7.08	21.25
3_3	13.63	7.83	
4_1	20.64	14.22	32
4_2	11.58	6.95	
4_3	17.65	13.22	
5_1	17.79	5.62	
5_2	18.67	8.23	
5_3	15.89	9.81	
6_1	35.83	24.69	
6_2	18.20	21.46	28
6_3	27.32	18.97	
7_1	8.17	4.88	28
7_2	9.64	5.45	
7_3	7.95	4.05	28
8_1	11.84	6.15	
8_2	9.84	4.59	28
8_3	11.73	5.99	26
9_1	17.22	9.16	24
9_2	13.64	5.50	18
9_3			
10_1	22.59	14.43	24
10_2	19.88	14.09	18
10_3			
11_1	39.83	32.48	24
11_2	3.30	0.04	
11_3			
12_1	24.41	17.80	30
12_2	22.91	7.14	28.5
12_3			

plot_id	avg_ba_m2/ha	avg_vol_m3	avg_anglcnt_m2/ha
1	23.24	14.76	29.00
2	11.83	6.07	31.50
3	20.67	9.82	24.63
4	16.63	11.46	32.00
5	17.45	7.89	
6	27.12	21.71	28.00
7	8.58	4.79	28.00
8	11.14	5.58	27.00
9	10.29	4.88	22.8625
10	14.16	9.50	20.42857143
11	21.57	16.26	24
12	23.66	12.47	29.68
avg	**17.19**	**10.43**	**27.01**

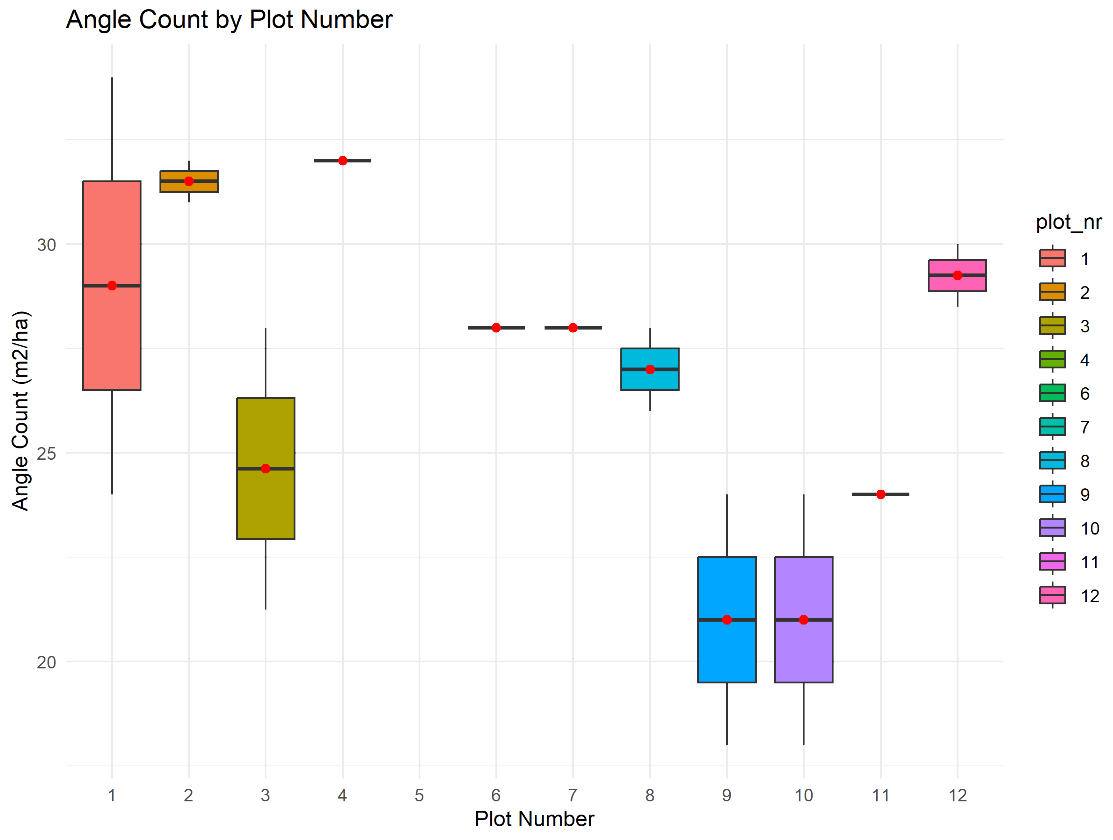
Results


Basal Area by Plot Number
Basal Area (m2/ha)
Plot Number
40
30
20
10
1
2
3
4
5
6
7
8
9
10
11
12
Plot Number
1 (mean: 23.24)
2 (mean: 11.83)
3 (mean: 20.67)
4 (mean: 16.62)
5 (mean: 17.45)
6 (mean: 27.12)
7 (mean: 8.59)
8 (mean: 11.14)
9 (mean: NA)
10 (mean: NA)
11 (mean: NA)
12 (mean: NA)

Results



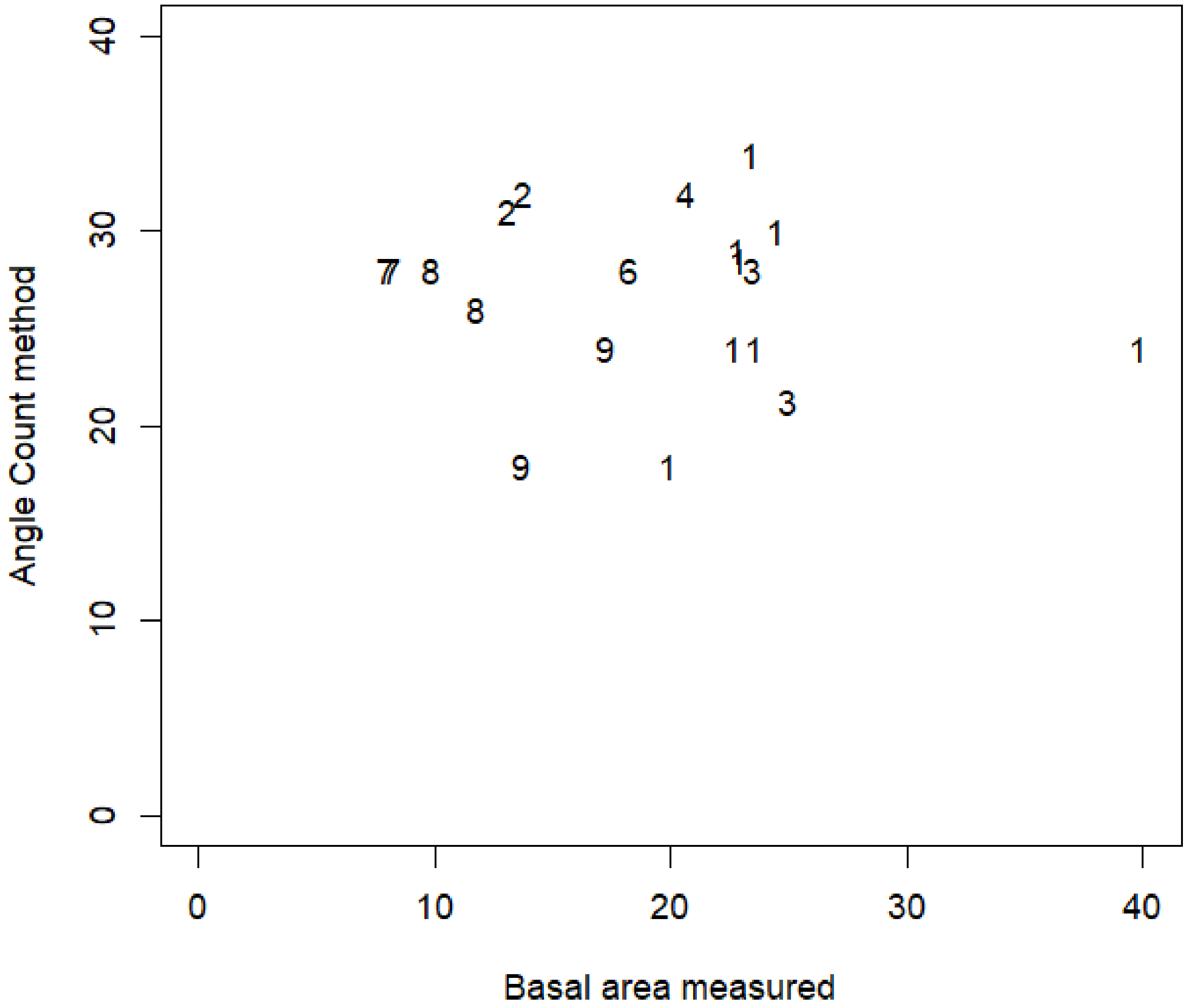
Basal Area by Plot Number
Basal Area (m2/ha)
40
30
20
10
1 2 3 4 5 6 7 8 9 10 11 12
Plot Number
Plot Number
1 (mean: 23.24)
2 (mean: 11.83)
3 (mean: 20.67)
4 (mean: 16.62)
5 (mean: 17.45)
6 (mean: 27.12)
7 (mean: 8.59)
8 (mean: 11.14)
9 (mean: NA)
10 (mean: NA)
11 (mean: NA)
12 (mean: NA)



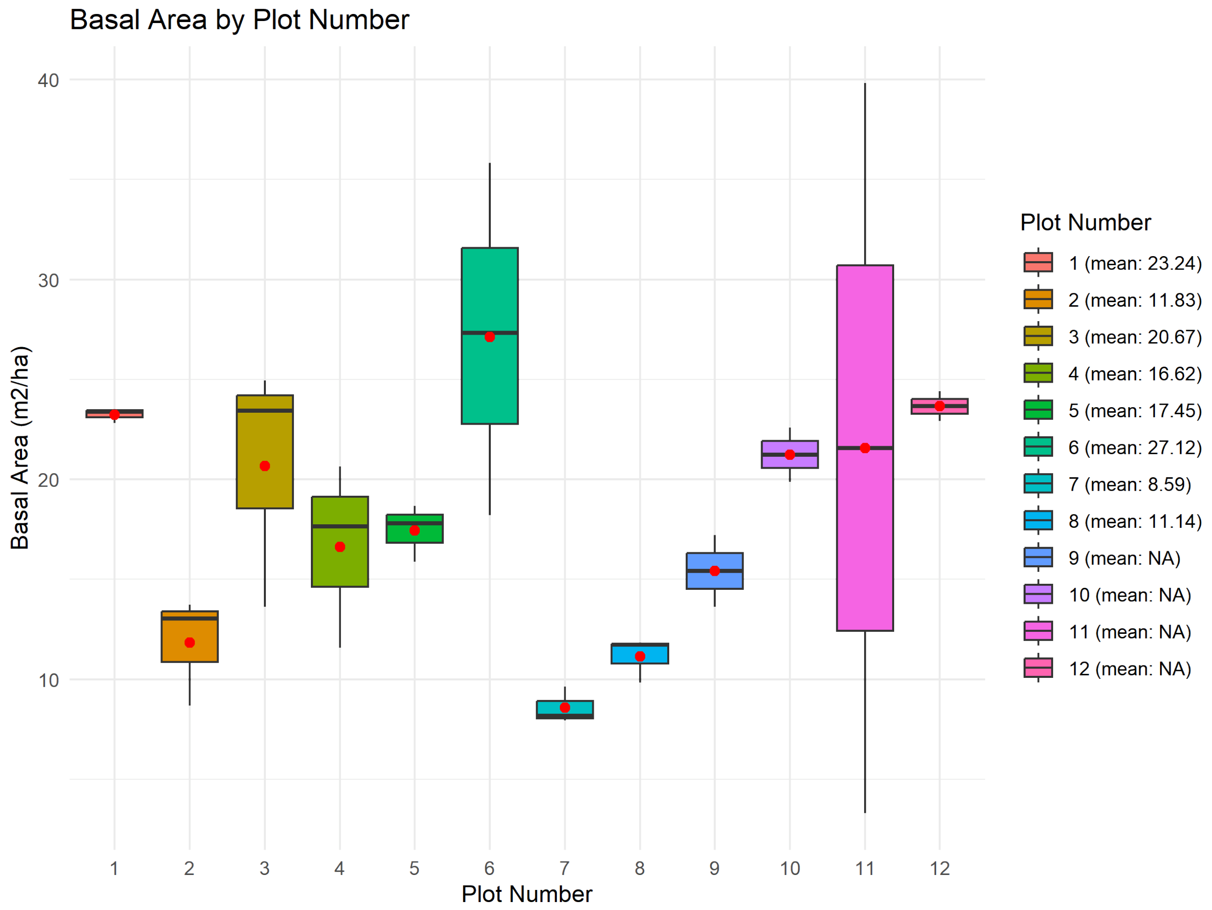
Angle Count by Plot Number
Angle Count (m2/ha)
30
25
20
1 2 3 4 5 6 7 8 9 10 11 12
Plot Number
plot_nr
1
2
3
4
6
7
8
9
10
11
12

Results

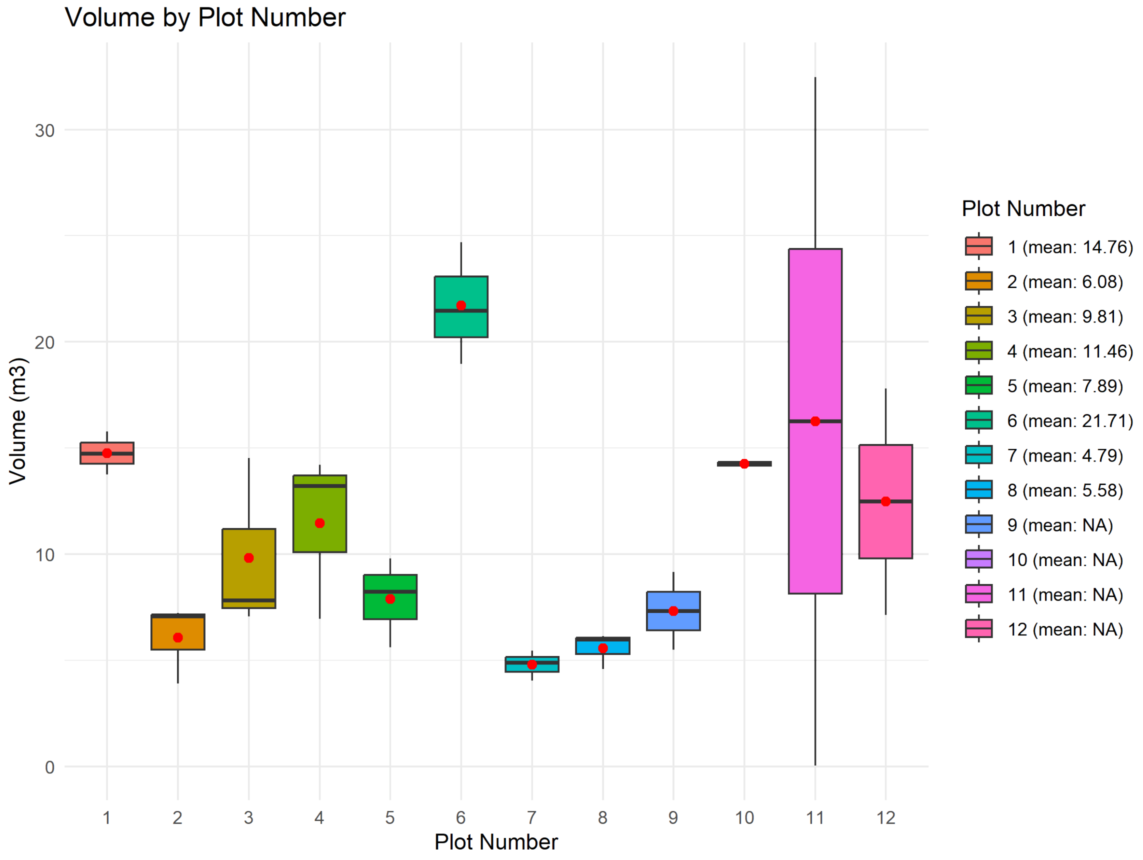
- Are Angle Counts and Basal area measurements correlated?
- Pearson´s r-squared: -0.12
- P-value: 0.6088



Results



Basal Area by Plot Number
40
30
20
10
Basal Area (m2/ha)
1
2
3
4
5
6
7
8
9
10
11
12
Plot Number
Plot Number
1 (mean: 23.24)
2 (mean: 11.83)
3 (mean: 20.67)
4 (mean: 16.62)
5 (mean: 17.45)
6 (mean: 27.12)
7 (mean: 8.59)
8 (mean: 11.14)
9 (mean: NA)
10 (mean: NA)
11 (mean: NA)
12 (mean: NA)



Volume by Plot Number
30
20
10
0
Volume (m3)
1
2
3
4
5
6
7
8
9
10
11
12
Plot Number
Plot Number
1 (mean: 14.76)
2 (mean: 6.08)
3 (mean: 9.81)
4 (mean: 11.46)
5 (mean: 7.89)
6 (mean: 21.71)
7 (mean: 4.79)
8 (mean: 5.58)
9 (mean: NA)
10 (mean: NA)
11 (mean: NA)
12 (mean: NA)

Results

Results

We could also look at the distribution of species, but the data was too inconsistent and many were missing.
We will continue to look at this in the exercise.

Species [41]
Carpinus betul
Quercus rubra
Quercus rubra
Alnus glutosa
Quercus rubra
Quercus rubra
Quercus rubra
Carpinus betul
Quercus rubra
Quercus rubra
Quercus rubra

Species [41]
HB
HB
RO
TP
HB
HB

Results

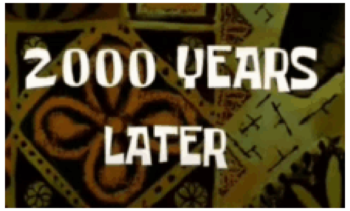
Summary

- There are many uncertainties and sources of variability
  - The human factor plays a very significant role (who measures how accurately)
  - Plot selection has an extremely high influence
  - Different measuring instruments have different accuracies
  - Technqiues also differ (actual survey vs. Angle count)

Results

Summary

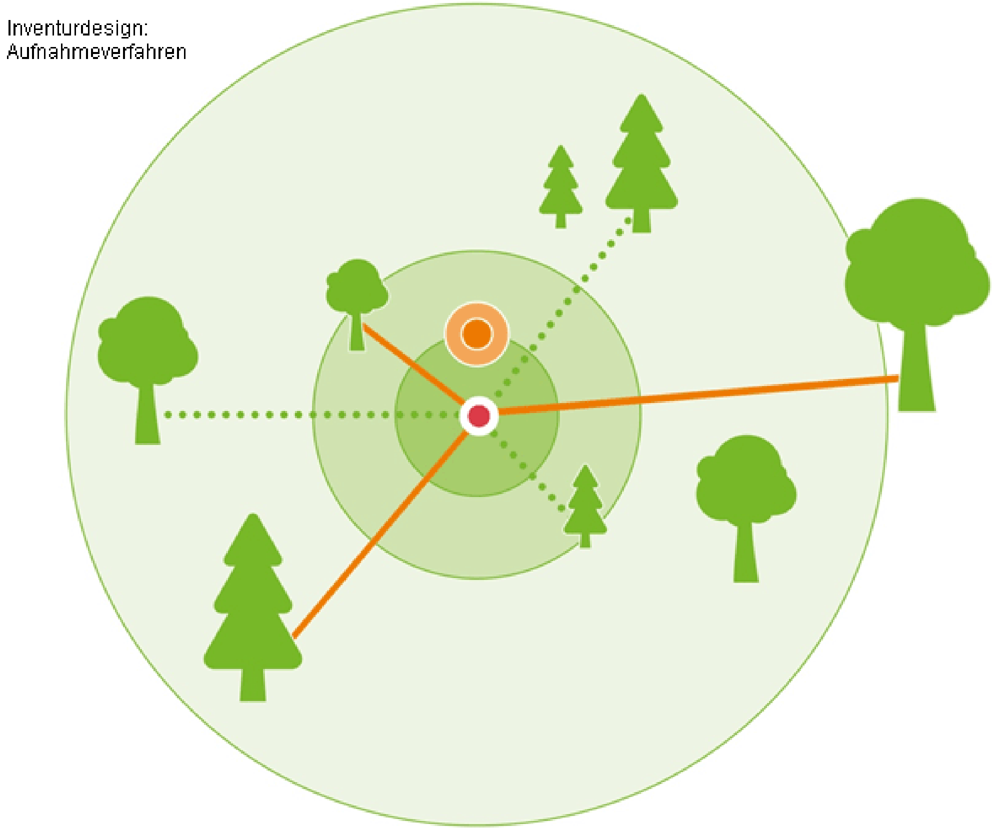
- There are many uncertainties and sources of variability
  - The human factor plays a very significant role (who measures how accurately)
  - Plot selection has an extremely high influence
  - Different measuring instruments have different accuracies
  - Technqiues also differ (actual survey vs. Angle count)
- …and it takes forever



Introduction to sampling uncertainty

definition

- Why don’t we record the entire tree population?

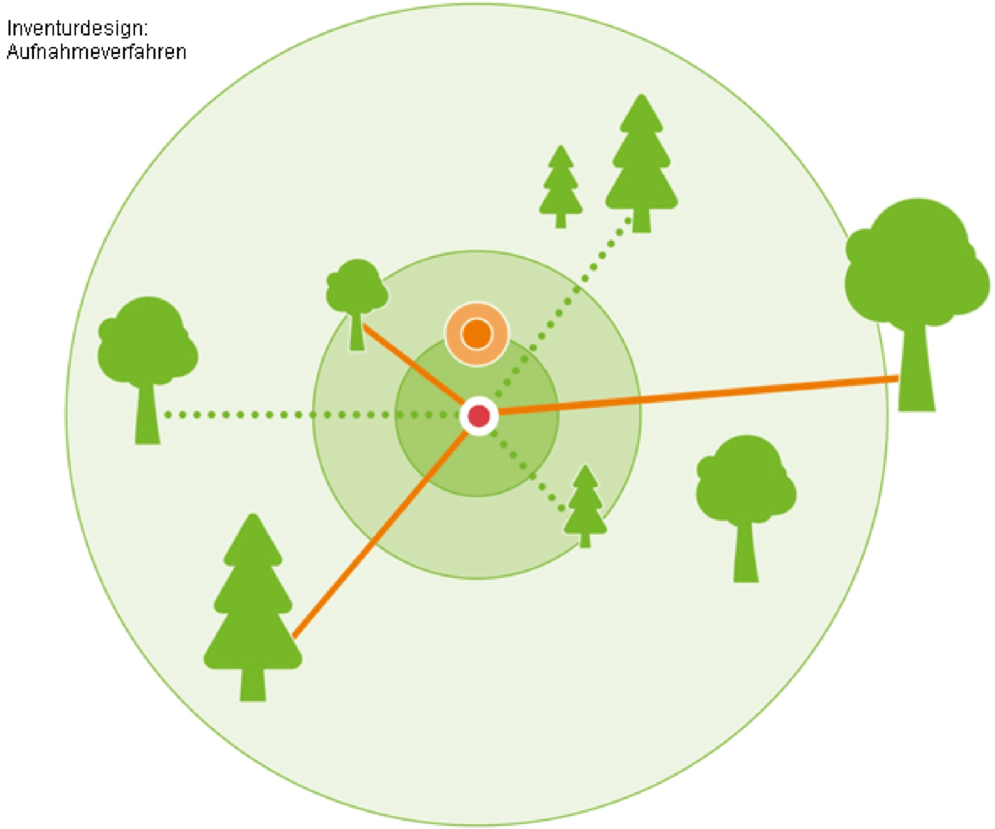


https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.waldwissen.net%2Fde%2Ftechnik-und-planung%2Fwaldinventur%2Feinfuehrung-in-die-bwi3&psig=AOvVaw3cej_rHtZPczaRHe_6ljko&ust=1731404386150000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCOjM7Iz-04kDFQAAAAAdAAAAABAE

definition

- Why don’t we record the entire tree population?

It would simply be too time-consuming and expensive



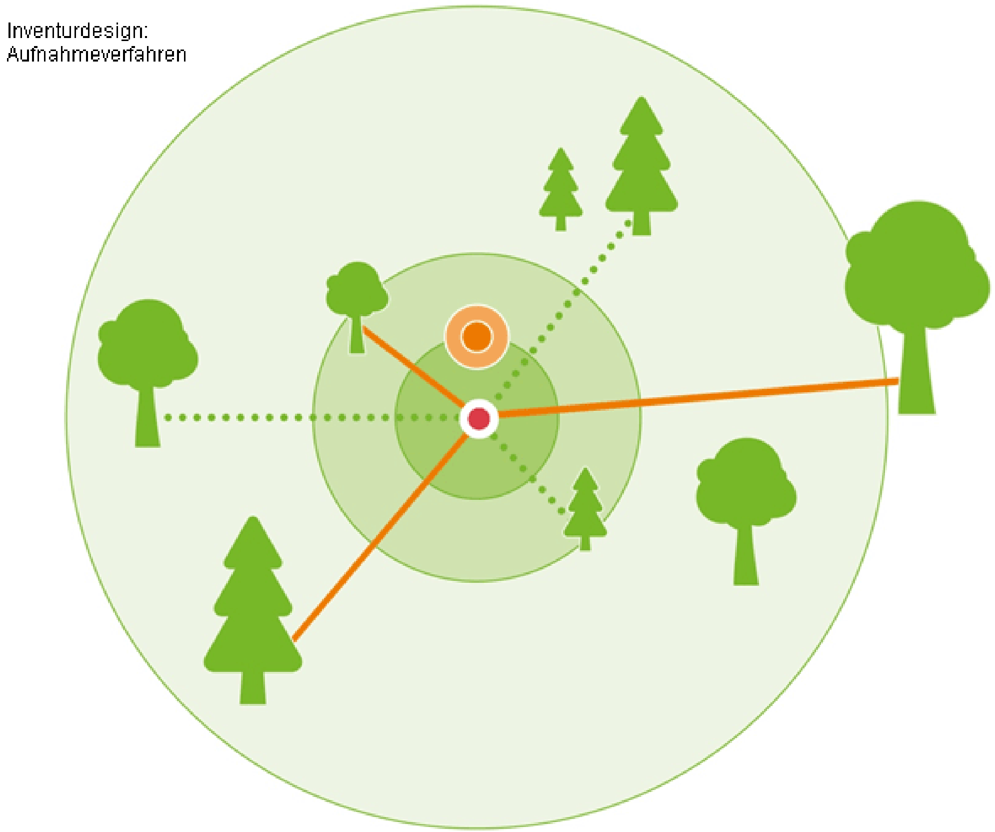
https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.waldwissen.net%2Fde%2Ftechnik-und-planung%2Fwaldinventur%2Feinfuehrung-in-die-bwi3&psig=AOvVaw3cej_rHtZPczaRHe_6ljko&ust=1731404386150000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCOjM7Iz-04kDFQAAAAAdAAAAABAE

definition

- **Sampling uncertainty** describes the uncertainties that result from using samples instead of full surveys.

  → what we did in the field exercise

- Relevance for the forest inventory: **A full inventory** is almost impossible or very time-consuming…
  In operational forestry, we have to rely on samples.

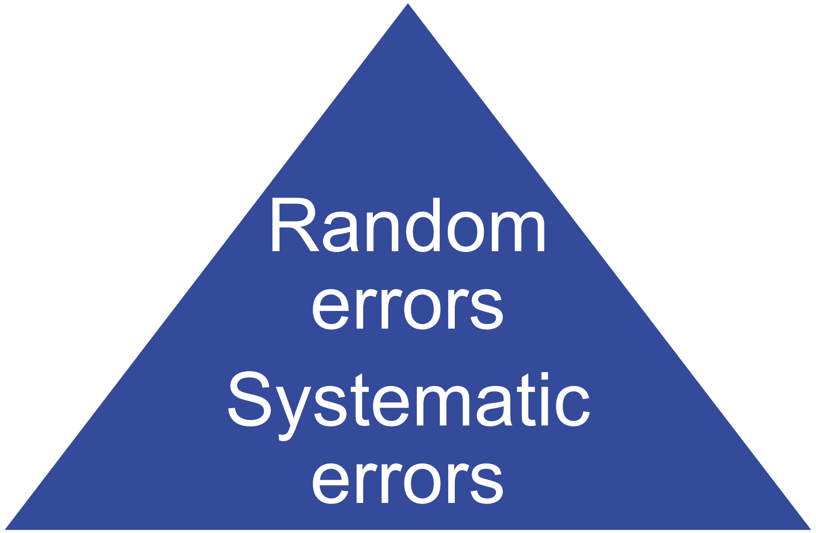


https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.waldwissen.net%2Fde%2Ftechnik-und-planung%2Fwaldinventur%2Feinfuehrung-in-die-bwi3&psig=AOvVaw3cej_rHtZPczaRHe_6ljko&ust=1731404386150000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCOjM7Iz-04kDFQAAAAAdAAAAABAE

uncertainties

Random errors:

- Unavoidable deviations, such as natural variability within the forest stand. Even if a random sample is taken, the selected area can differ greatly in its properties from the average of the entire stand, leading to inaccuracies.
- Random errors from measurements



´Systematic errors:

- Errors resulting from inadequate or distorted selection methods.
- Faulty instruments or incorrect methods
- Inadequate or faulty selection methods (e.g. biased plot selection)
- Accessibility
- subjective influences (e.g. human bias)

quantification of uncertainty

- **Confidence intervals**: Range within which the true value (e.g. a mean) of the entire population lies with a certain probability (e.g. 95%).
  - Larger samples lead to narrower confidence intervals and thus to greater precision – but more work required.

- **Standard error (SE):** Measure of the accuracy of the estimate to the true populatiuon parameter (e.g. a mean basal area).

- **Variance (S²)**: Spread of the individual data points around the mean (e.g. the spread of the basal area of the individual samples to the mean of all samples)

$$CI = \bar{x} \pm z \frac{s}{\sqrt{n}}$$

CI = confidence interval
$\bar{x}$ = sample mean
z = confidence level value
s = sample standard deviation
n = sample size

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

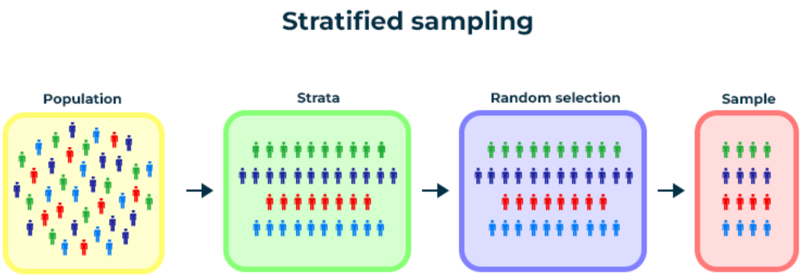
σ = population standard deviation
N = the size of the population
x_i = each value from the population
μ = the population mean

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

S^2 = sample variance
x_i = the value of the one observation
$\bar{x}$ = the mean value of all observations
n = the number of observations

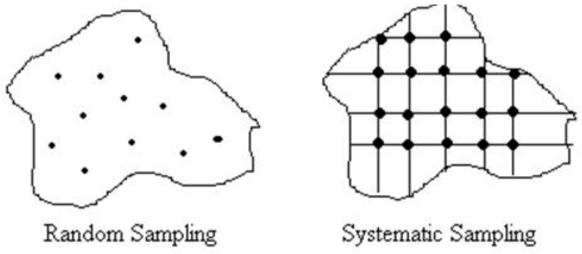
Reducing uncertainty

- In general: Increase sample size - More measurement points improve accuracy and reduce uncertainty.



https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.geeksforgeeks.org%2Fstratified-random-sampling-an-overview%2F&psig=AOvVaw0-THJ857kquPKlAo4HLk5t&ust=1731586841705000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCJjrreal2YkDFQAAAAAdAAAAABA4

1) **Stratified sampling**: The forest is divided into homogeneous layers (e.g. by tree species or age) and separate samples are taken for each layer. This increases precision because differences between layers are taken into account.



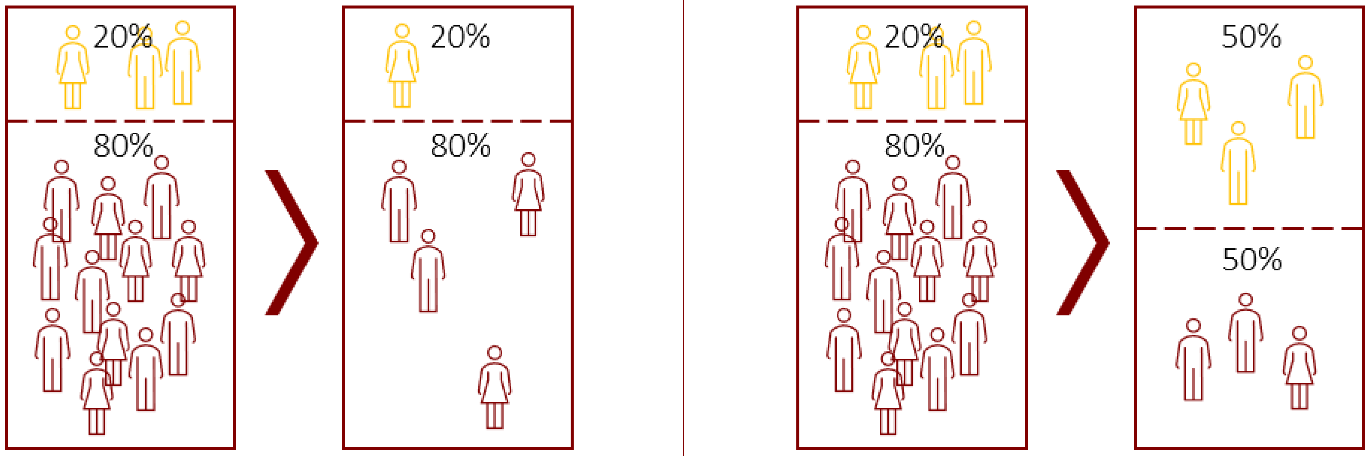
https://www.google.com/url?sa=i&url=https%3A%2F%2Fforestrypedia.com%2Fthe-forest-resource%2F&psig=AOvVaw2iWvngqhGBn08rVFWvlw_w&ust=1731586906681000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCMCLpYmm2YkDFQAAAAAdAAAAABAE

2) **Systematic sampling**: Instead of randomly, samples are taken at fixed intervals, e.g. every 100 meters along a grid. This method can ensure better coverage and more even distribution.

reducing uncertainty

1) Example forest:

- Stock 80% beech, 20% spruce
- → If I were to represent 50% spruce in my sample, I would have a distorted image of the entire forest stock.



Whole data

Proportional sample

Whole data

disproportional sample

Summary

- We now know
- ... what sampling uncertainty is
- ... why it is important for forest inventories
- ... that there are random and systematic errors
- ... how we can quantify these errors
- ... how we can address the uncertainties

Summary

- We don't yet know
- ... how large these uncertainties are in a large, real forest stand

- → Exercise in R

Questions?

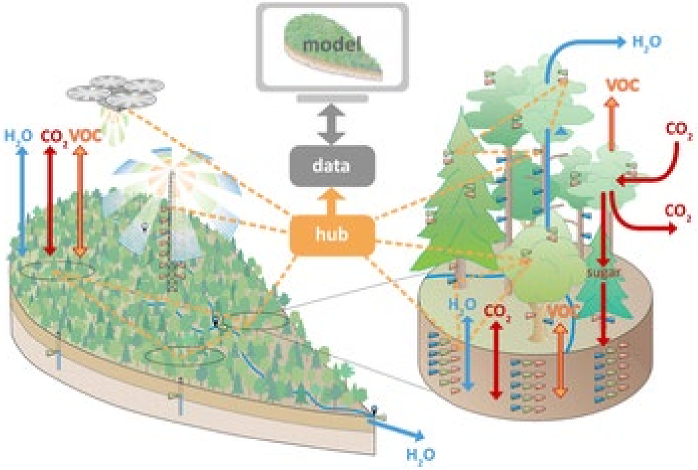
- What questions do you have so far?

Transition to the exercise in R

We will now simulate one or more samples.

Full inventory of Ecosense
https://ecosense.uni-freiburg.de/

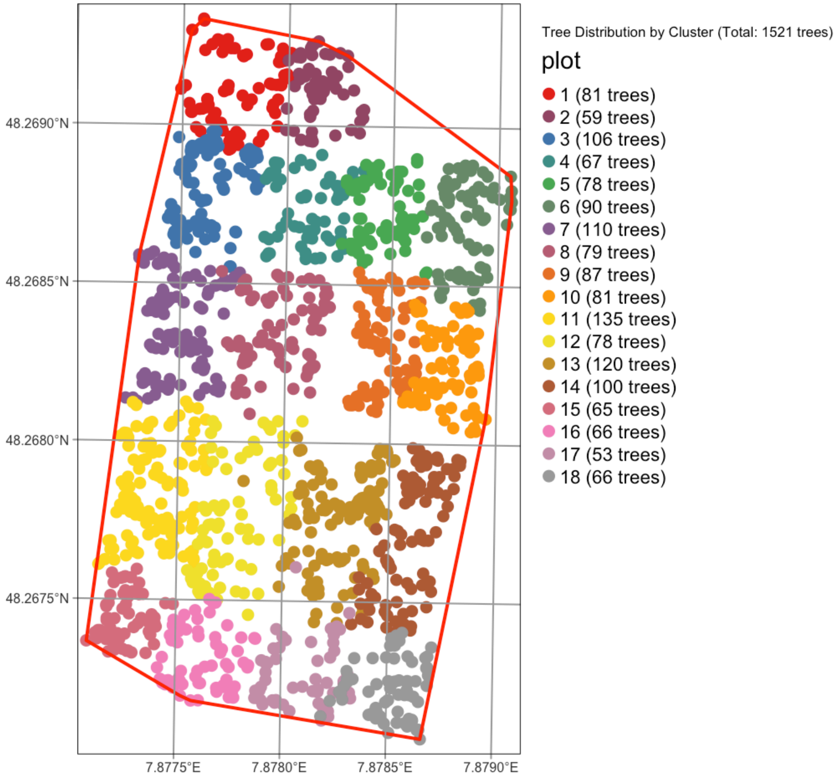
- **Research site** near Ettenheim
- ECOSENSE researches critical ecosystem changes and develops an intelligent sensor network.



Why relevant for us? We know each tree on this site on the basis of a full inventory.

Full inventory of Ecosense https://ecosense.uni-freiburg.de/

- Download the file ex05-pw.zip and unpack the data where you can find it again.
- Password is: Inventory
- Open the .qmd file in Rstudio
- The geopackage contains tree coordinates, as well as many other attributes of 1521 trees



What exactly are we doing today

Basic statistics

Simulate sampling

- random sampling and regular grid

“real scenario”

- Influence of GNSS uncertainty to Plot measurements

Small refresher of the required packages in R

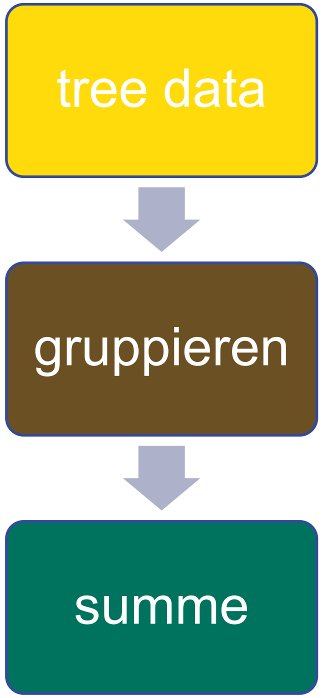
Working with geodata → General feature handling - buffer, clip etc.

Working with geodata → Nice tool for creating maps

Package	Description
`sf`	Provides simple features for handling and analyzing spatial data, making it easy to work with geometries.
`dplyr`	A grammar of data manipulation that helps in data wrangling tasks such as filtering, selecting, and summarizing.
`tmap`	A package for thematic mapping, offering an easy way to create static and interactive maps.
`ggplot2`	A powerful system for creating complex and layered graphics using the grammar of graphics.
`units`	Provides support for handling measurement units in R, ensuring unit-aware arithmetic and conversion.
`reshape2`	Helps in transforming and reshaping data, particularly useful for converting between wide and long formats.

dplyr

```
# Calculate number of trees per cluster/plot
trees_per_plot <- tree_inventory_data %>%
                  group_by(plot) %>%
                  summarize(tree_count = n())
```



This section of code uses the %>% operator

The operator enables what is known as pipelining, a concept in which the output of one function is passed directly as input to the next function.

sf package

```
# Read a specific layer from the GeoPackage
tree_inventory_data <- st_read("tree_inv_ike.gpkg", layer =
"datasheet_ecosense2023")
```

The sf package in R is used for working with spatial (geographical) data.

Geodata analysis
based on the Simple Features Standard

sf package

1. Import and export data
st_read() and st_write(): Imports/exports spatial data in various formats, e.g. shapefiles, GeoJSON, KML.

2. Create and manipulate geometries
st_point(), st_linestring(), st_polygon(): Creates points, lines and polygons.
st_geometry(): Extracts or sets the geometry column in an sf object.
st_union(), st_intersection(), st_difference(): Performs geometry operations such as union, intersection or difference.

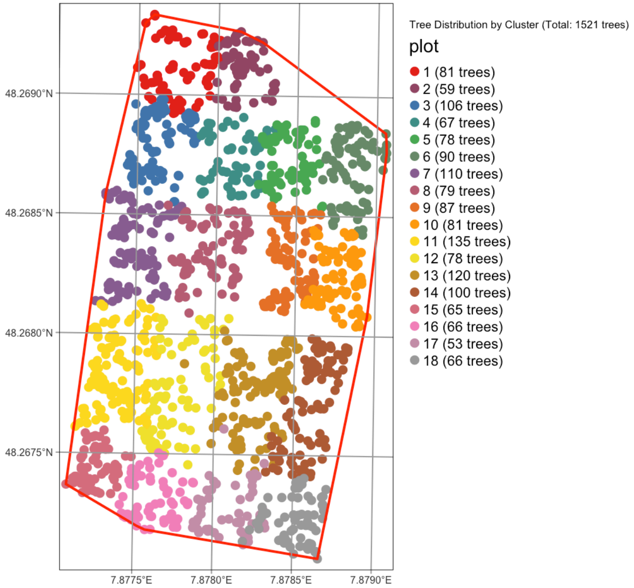
3. Spatial calculations
st_area(), st_length(), st_distance(): Calculates area, length and distance of geometric objects.
st_centroid(): Calculates the center of polygons.

4. Transformations and projections
st_transform(): Changes the projection of a spatial object (e.g. from geographic coordinates to a Cartesian projection).
st_crs(): Sets or returns the coordinate system.

5. Processing and analyzing data
st_join(): Combines two sf objects based on spatial relationships (e.g. within, overlapping).
st_buffer(): Creates a buffer around geometries.

tmap package

```
# Plot using tmap with unique colors for each plot and updated legend
tm_shape(tree_inventory_data) +
  tm_dots(col = "plot", palette = "Set1", size = 0.2) + # Color points by 'plot' column
  tm_shape(field_extent) +
  tm_borders(col = "red", lwd = 2) + # Add extent as red borders
  tm_graticules(lines = TRUE, labels.size = 0.5, col = "gray60", n.x = 4, n.y = 4) +
  tm_layout(legend.outside = TRUE,  # Place the legend outside the plot
            title = paste("Tree Distribution by Cluster (Total:", total_trees, "trees)")) #
Set the title with total trees
```



tm_shape(): Sets the data basis for the map (e.g. point or area data), which is then visualized in further steps.

+ tm_dots(): Draws points for each tree position and colors them according to the "plot" value, so that each group gets a different color.

+ tm_borders(): Adds a border line around the shape data (e.g. the study area), here red and slightly thicker.

+ tm_graticules(): Adds a grid to the map, which creates orientation and reference points in the background.

+ tm_layout(): Adjusts the layout, e.g. placing the legend outside the map and setting a title with the total number of trees.

ggplot2 package

ggplot2 is a powerful and flexible R package for data visualization based on the concept of the so-called Grammar of Graphics.

→ Modularly customizable (similar to tmap)

Basic structure:

ggplot(data = <data>) +
aes(x = <X-axis variable>, y = <Y-axis variable>) +
geom_<type>() +
theme() +
labs()

Points, polygons, boxplot, etc.

fonts etc.

Labels/Text

Switch to Rstudio (if you want)

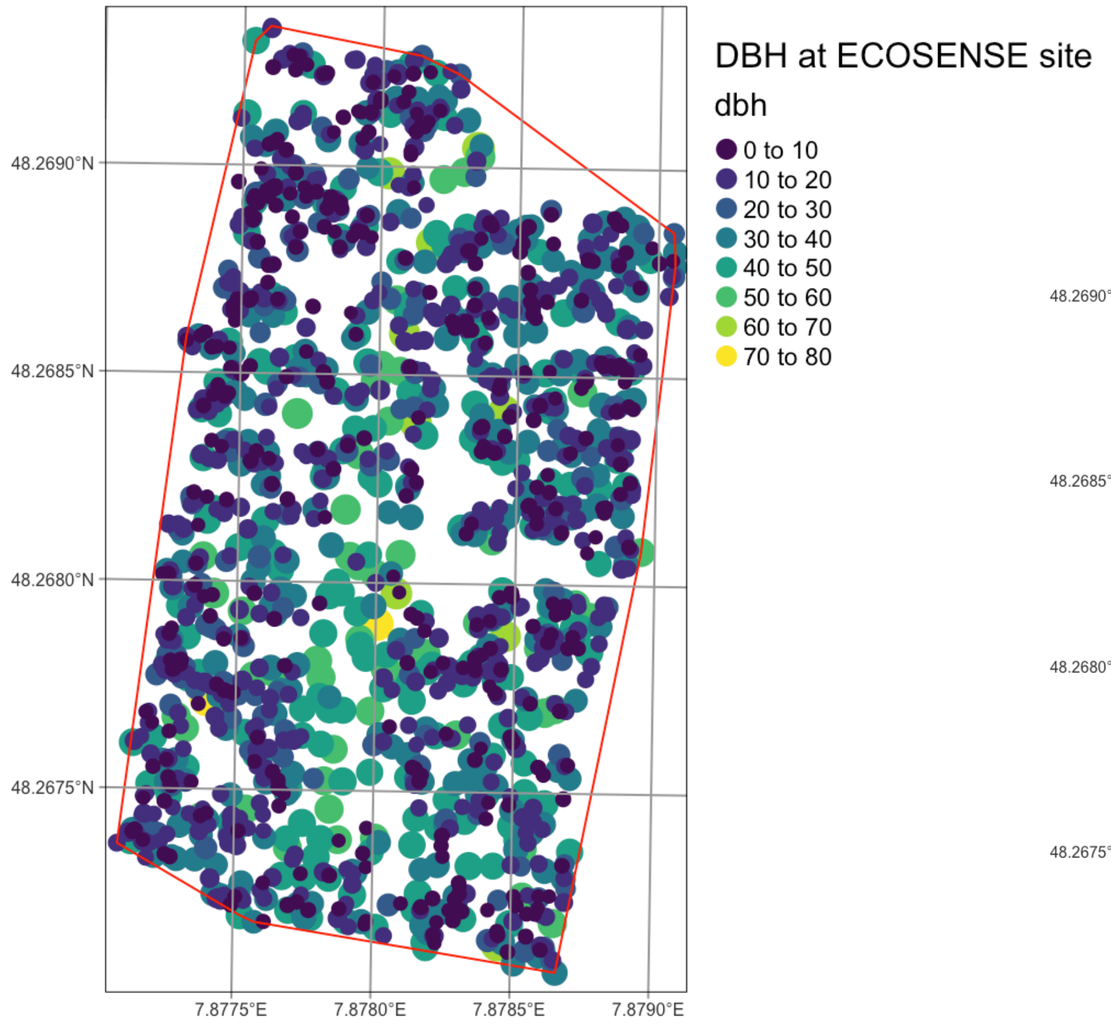
Now we will look at the exercise in Rstudio

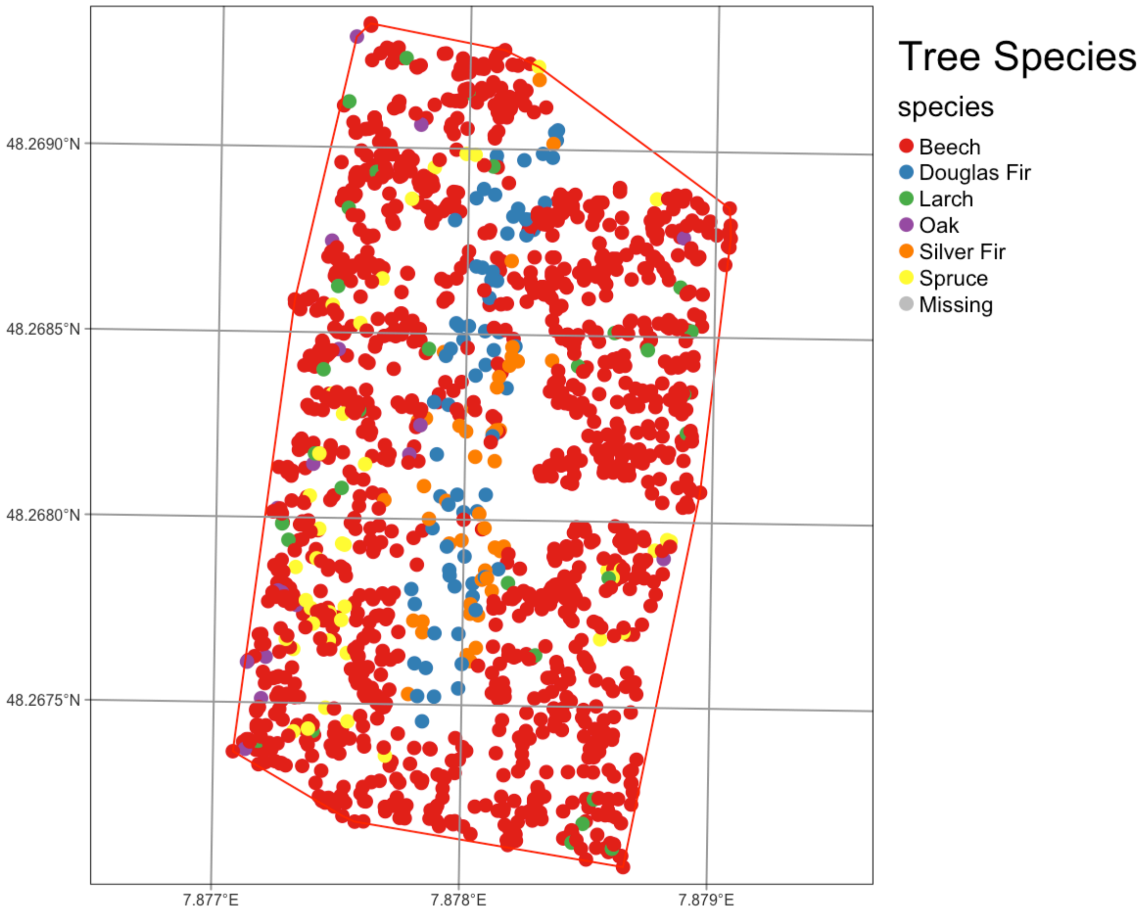
Basic structure of the document:

- Loading and viewing data, using sf and tmap packages
- Statistics about the forest cover (basal area, species, DBH)
- Plot selection and uncertainties
- Simulation
- Questions to answer (woohoo)

Viewing data in R

What do we see?

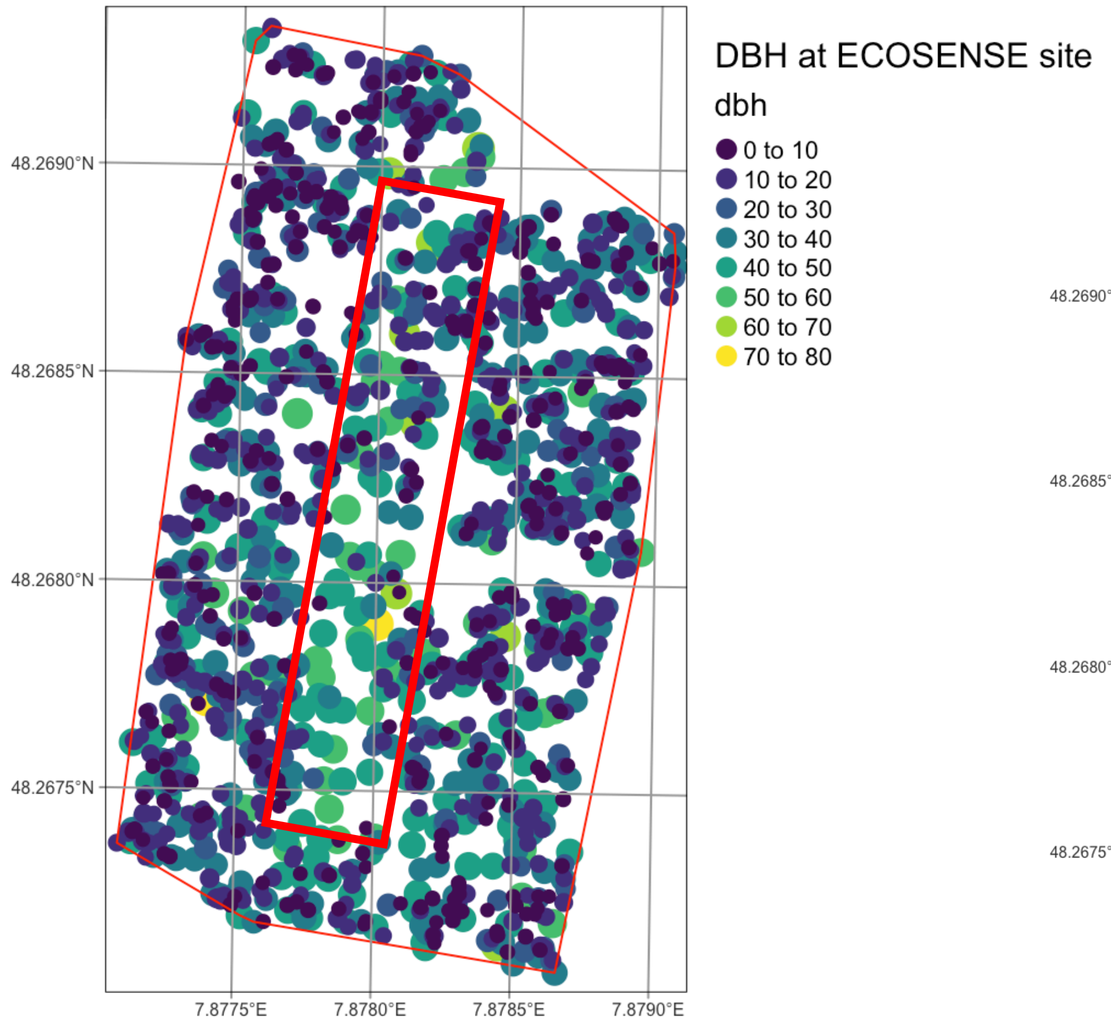


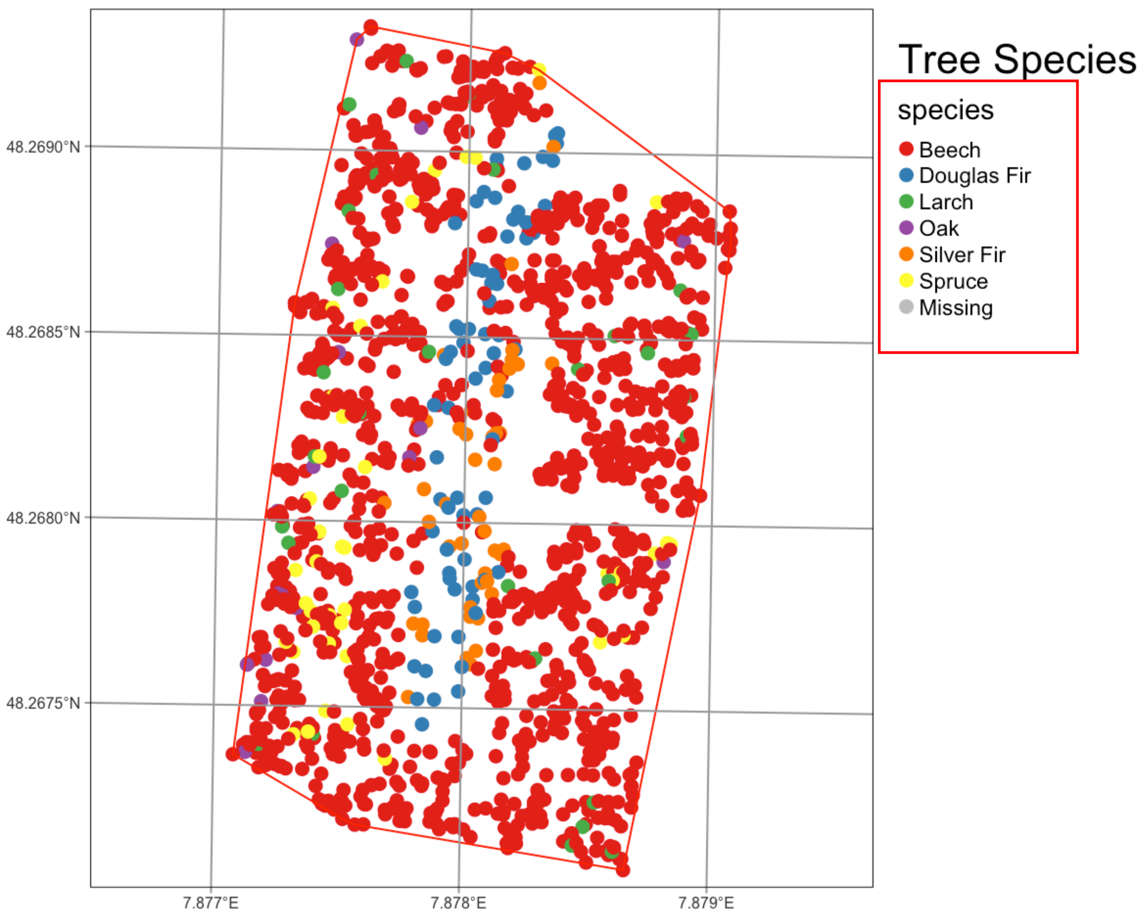


Viewing data in R

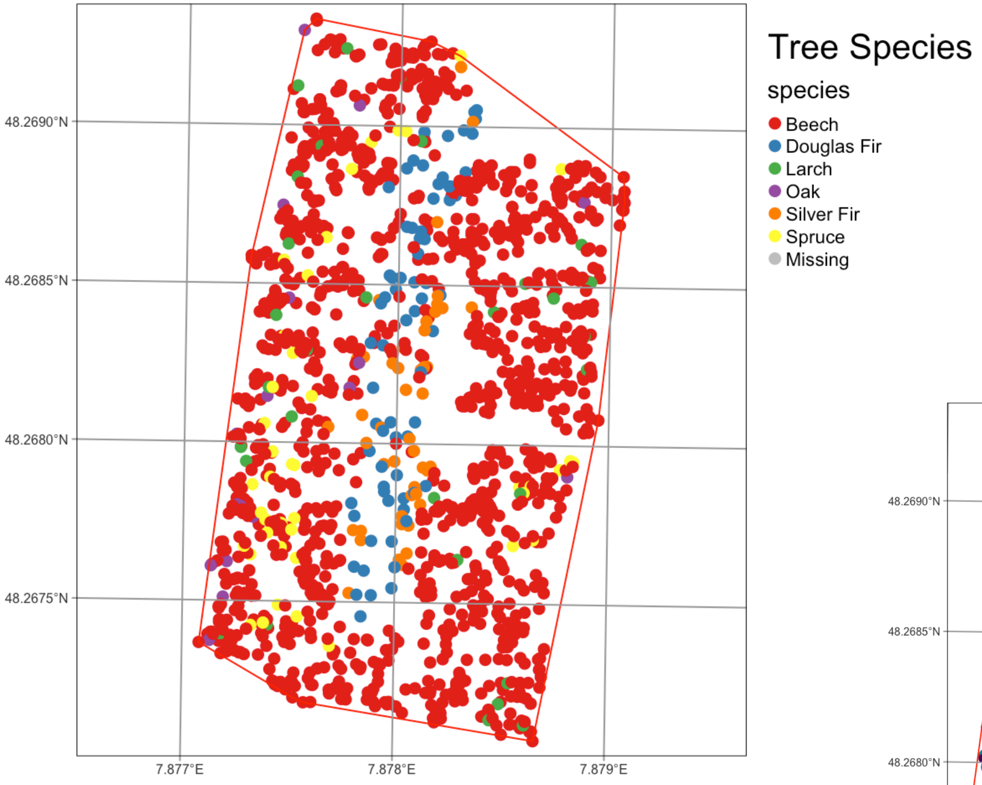
Right Small trees outside, large trees in the middle

Many beeches, many Douglas firs



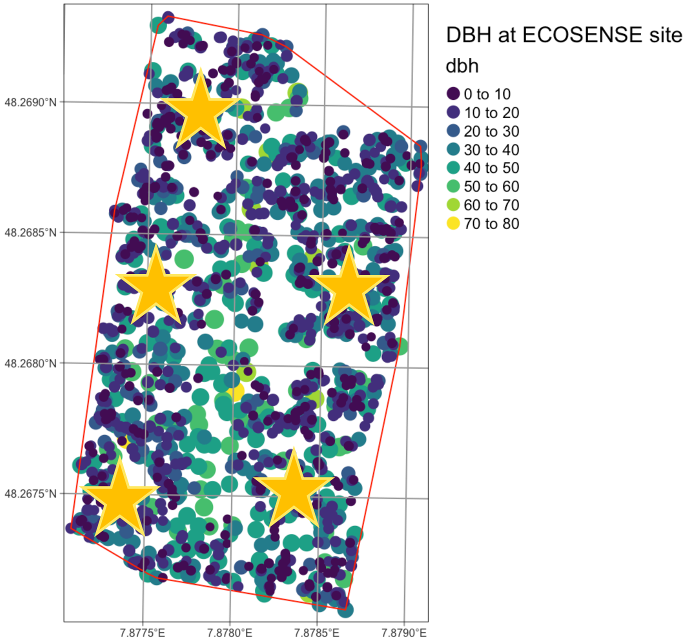


Viewing data in R

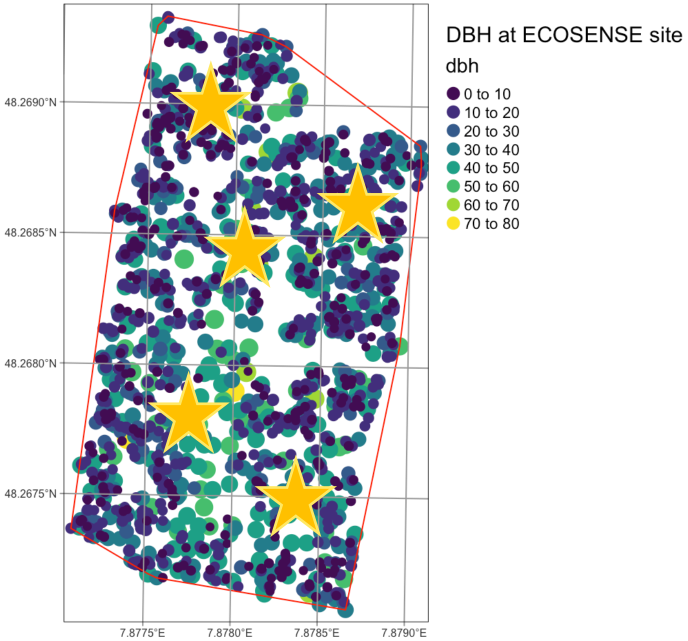


Think back to the exercise – which option would (probably) be the more meaningful:

1



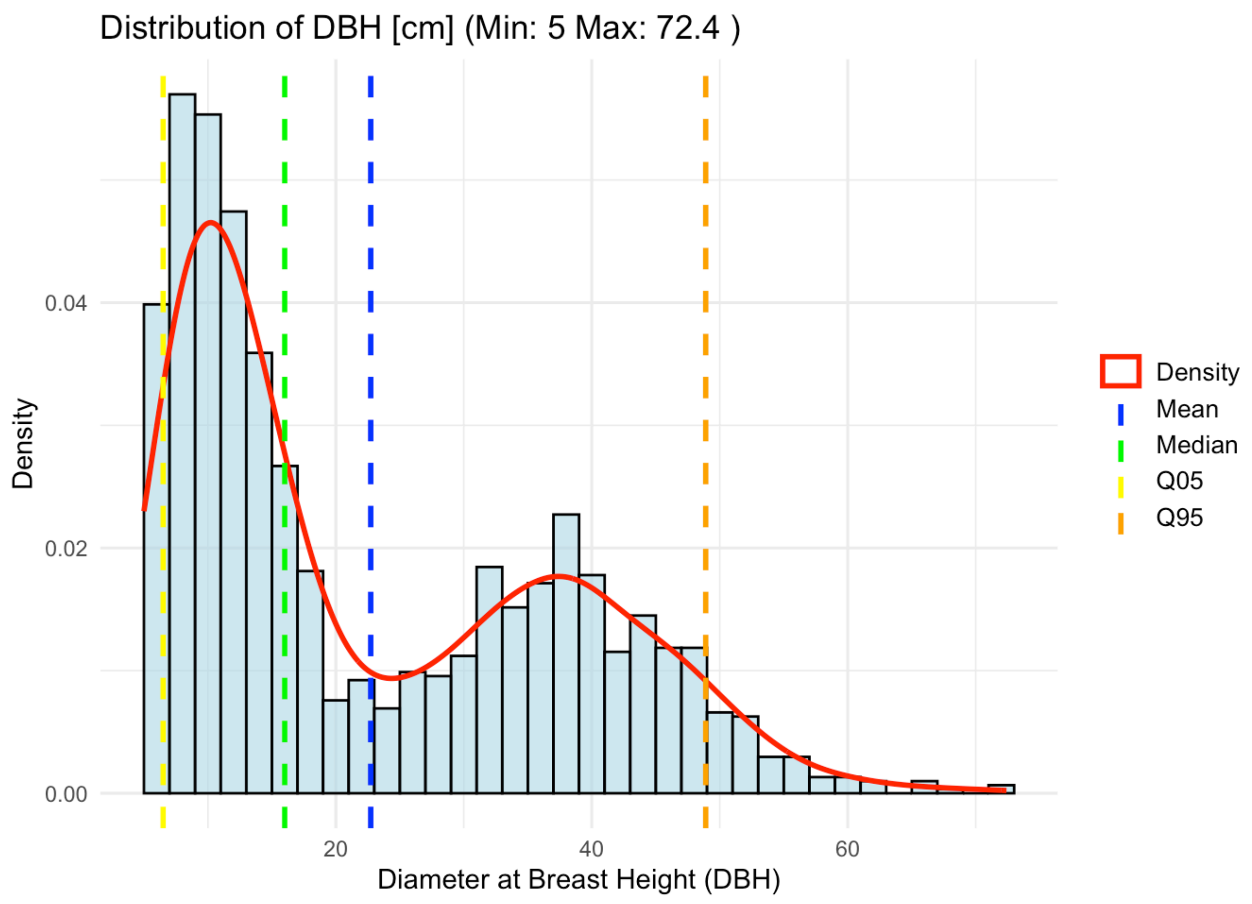
2



Viewing data in R

What does this graphic tell us?

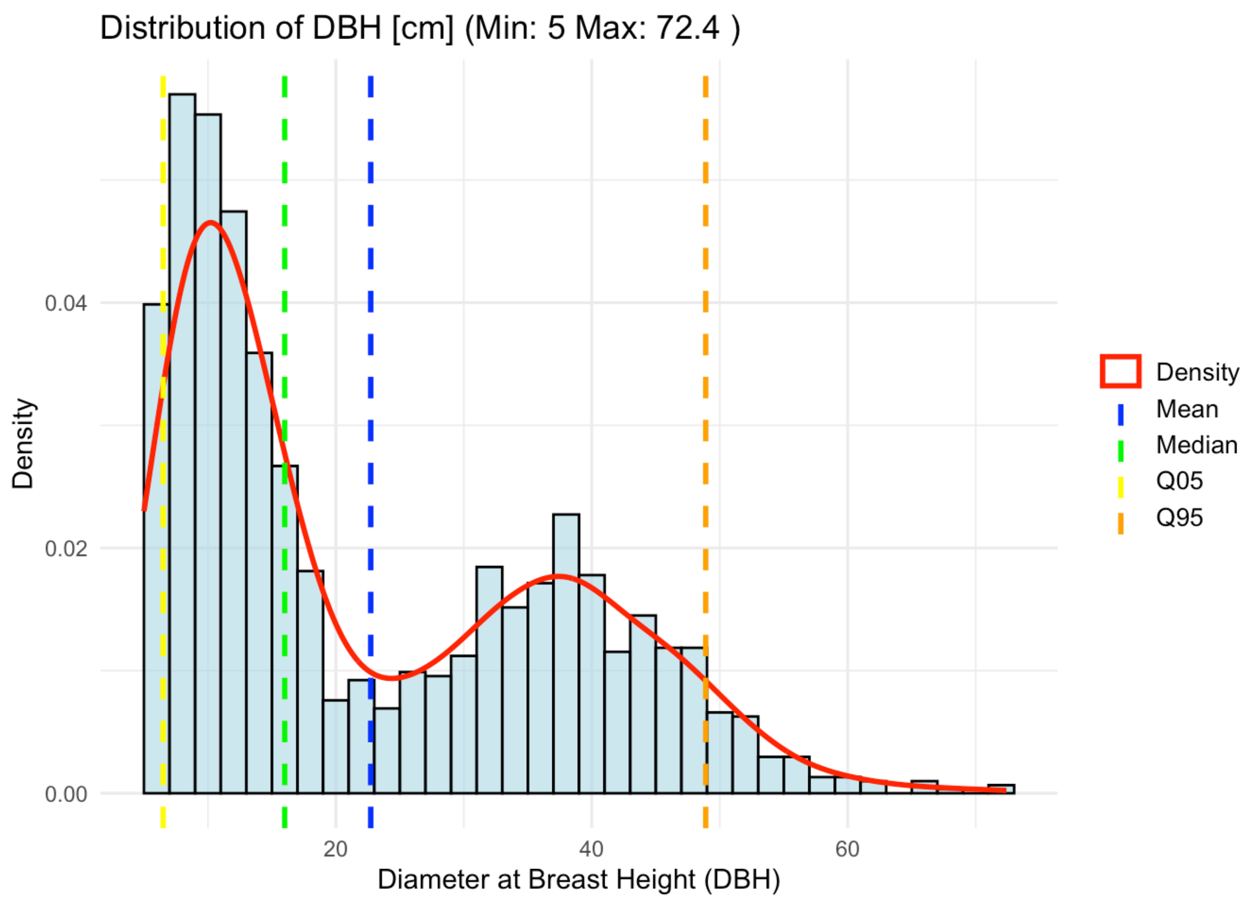
Small trees? Big trees? Well distributed? Young or old forest?



Viewing data in R

Bimodal distribution of trees
Hence high deviations from the average

Base area calculated from this: 31.68 m2/ha



Viewing data in R

Distribution and significance of DBH (cm):

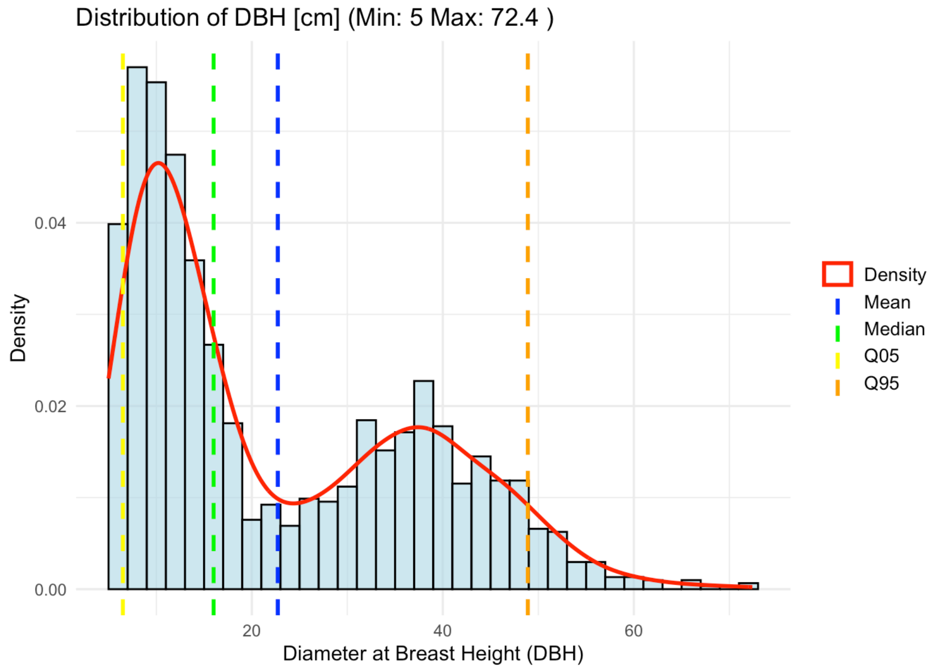
Mean (22.72) vs Median (16): Median lower than the mean, indicating right-skewed distribution.

Mode (9): Most frequent value, lower than mean and median, also supports right-skewed distribution.

Standard deviation (14.89): Measure of dispersion around the mean; high value indicates variability.

IQR (25.9): Range of the middle 50% of the data; large value compared to the median indicates wide range of values.

Summary: The distribution is right-skewed (mean > median > mode) with high variability (high standard deviation and IQR).
→ A few very large values/trees

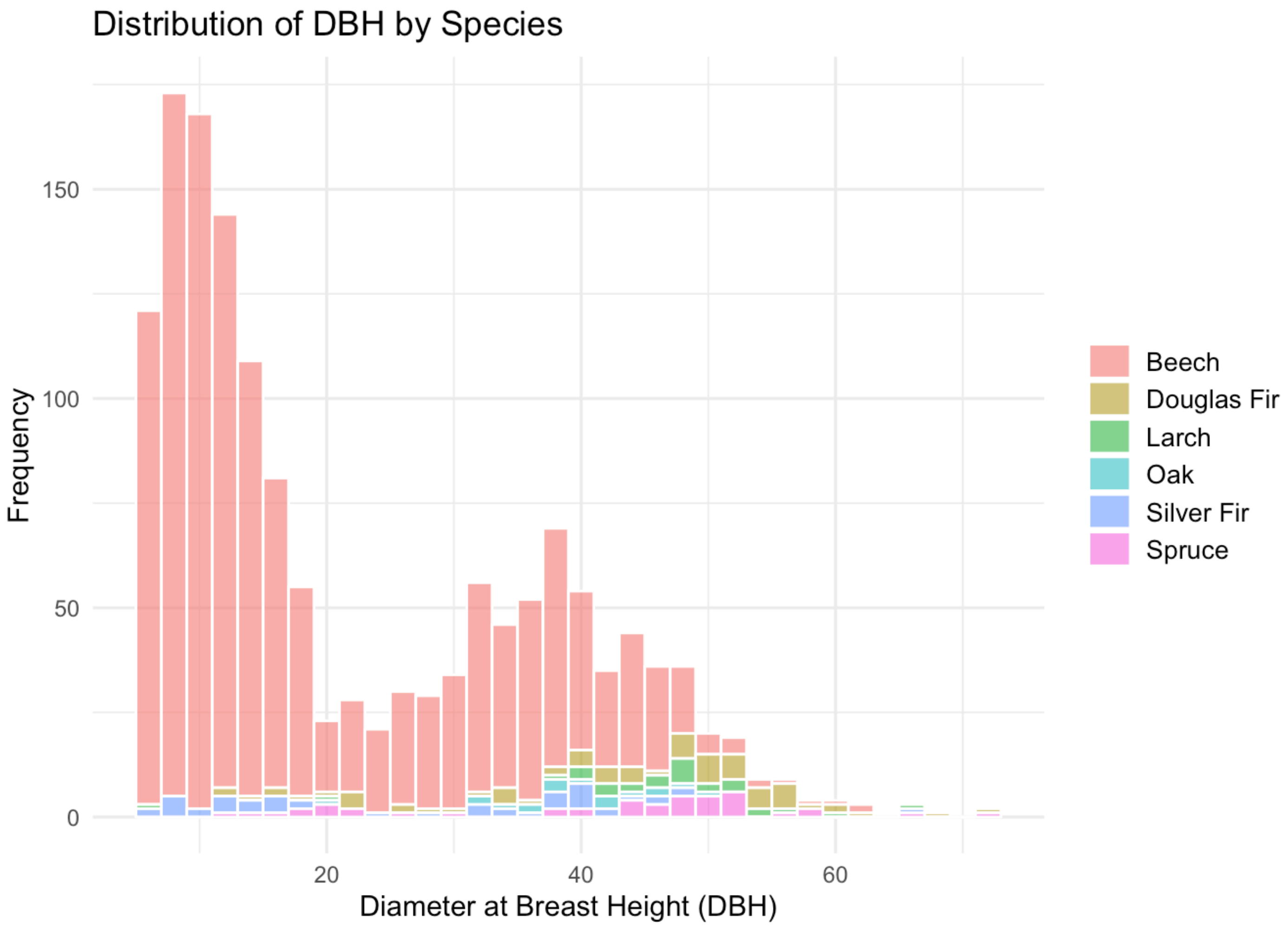


Viewing data in R

Species distribution

What does this graph show us?

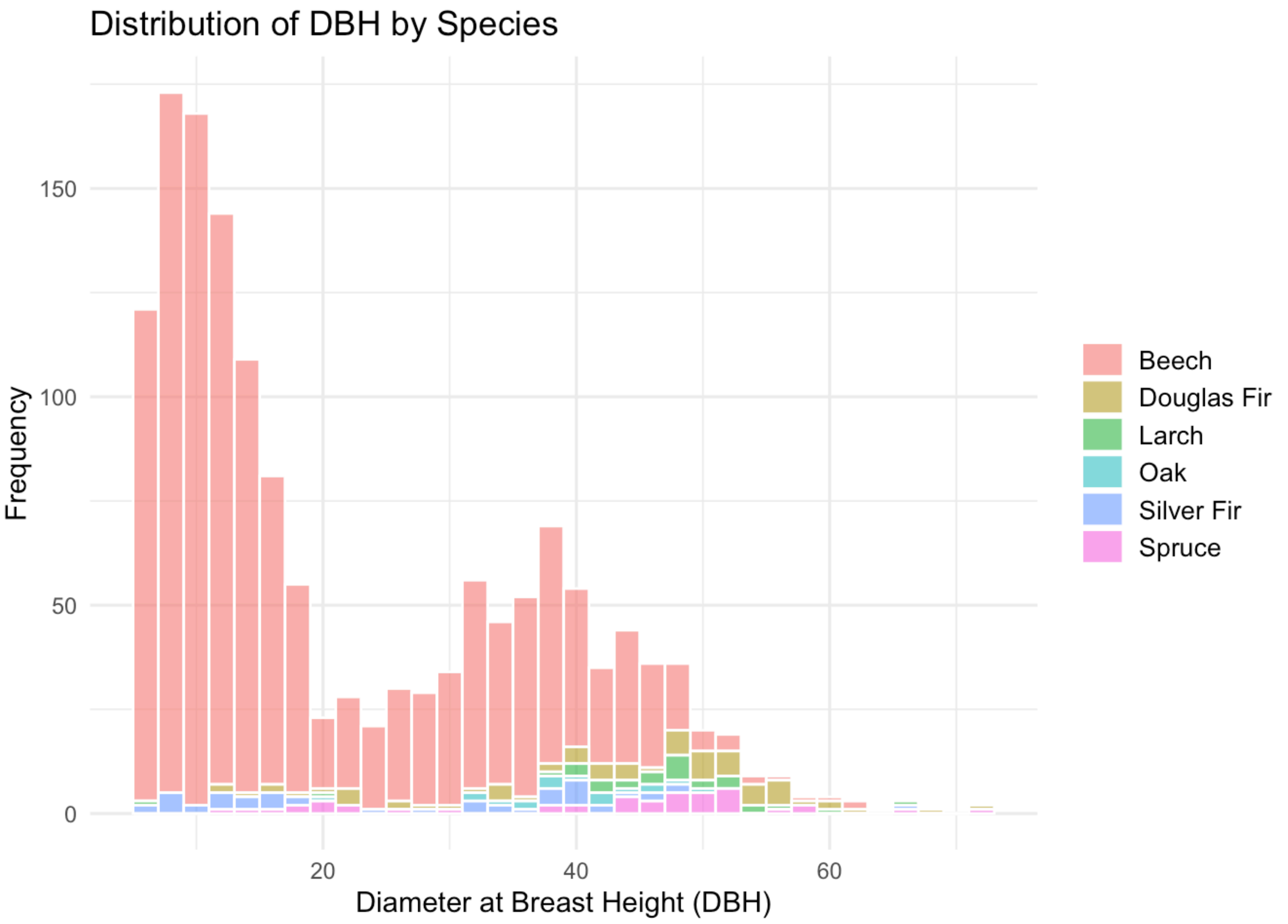
What do you notice?



Viewing data in R

Species distribution

Colors are relative to the entire stand
Many young beeches
Some larger oaks, Douglas firs and larches
…



Rest in the exercise in the afternoon

We will simulate one or more samples.

Viewing data in R Simulation of samples – Random sampling and regular grid

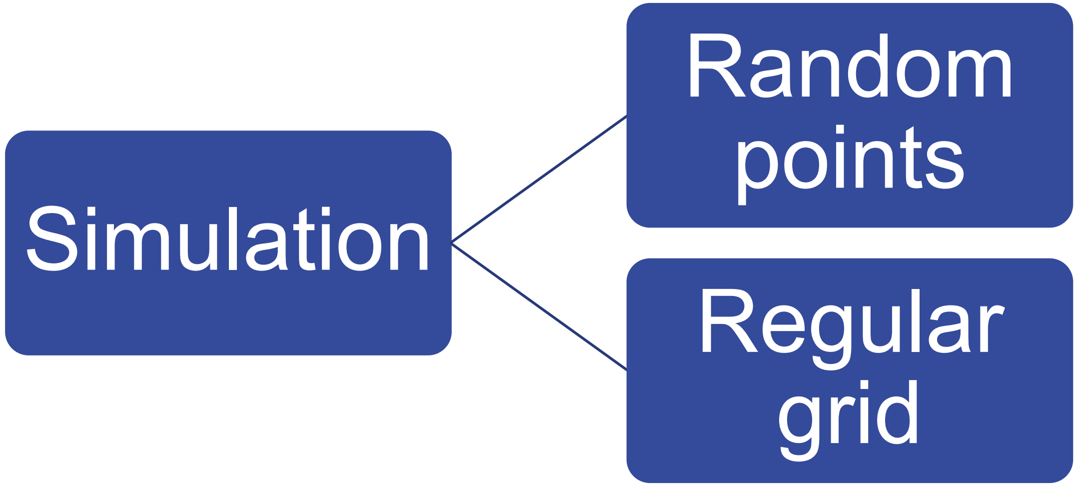
We do 2 simulations and see what happens:

1. Random plots
2. Regular grid

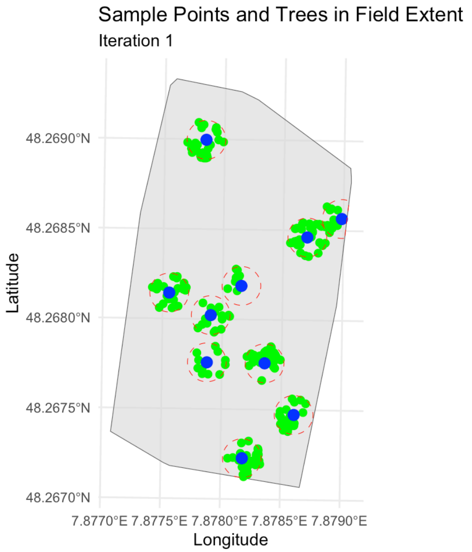
In the example, we take a sample of 10 randomly distributed points (plots) in the existing area 100 times

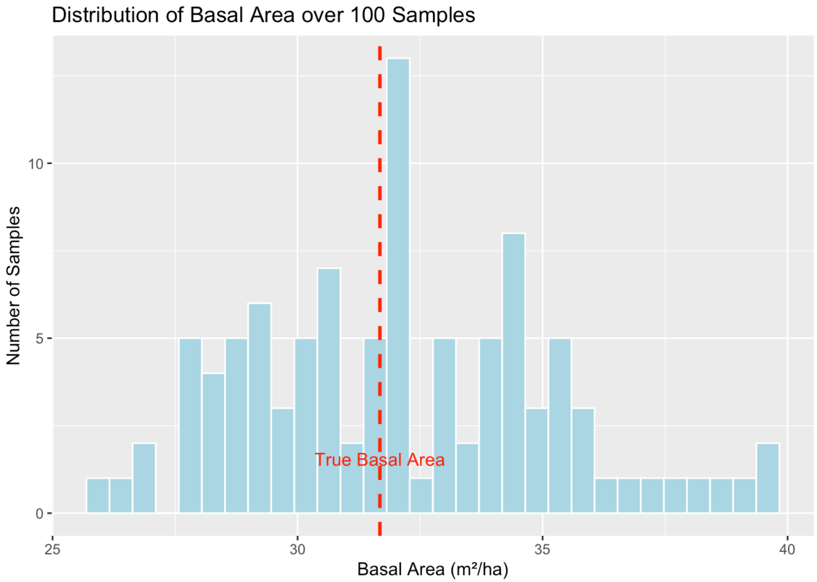
From this we can derive the basal area each time and see how well (or badly) we did

We do this for the 2 options of plot distribution mentioned and look at the results!



Viewing data in R Simulation of samples – Random sampling and regular grid

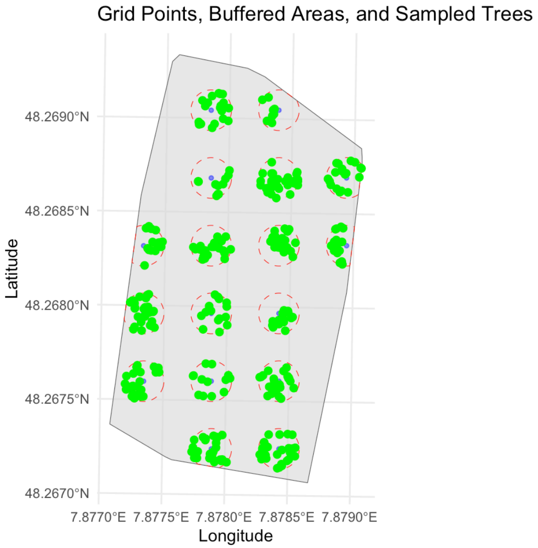


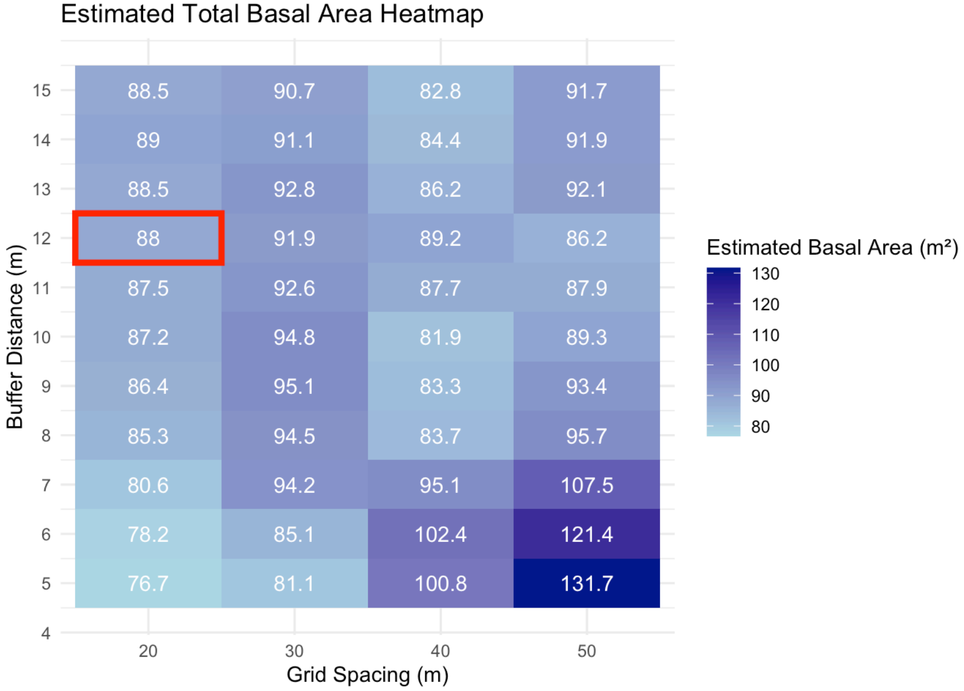


What do you notice? Is that good? Is that bad?

“that’s just how it is?”

Viewing data in R Simulation of samples – Random sampling and regular grid





A regular grid is a simple and good method of getting a realistic image. Don't stick to one value (here, for example, 20/12) - this varies depending on the area and size!
There are also some extremes. It also requires a lot of work!

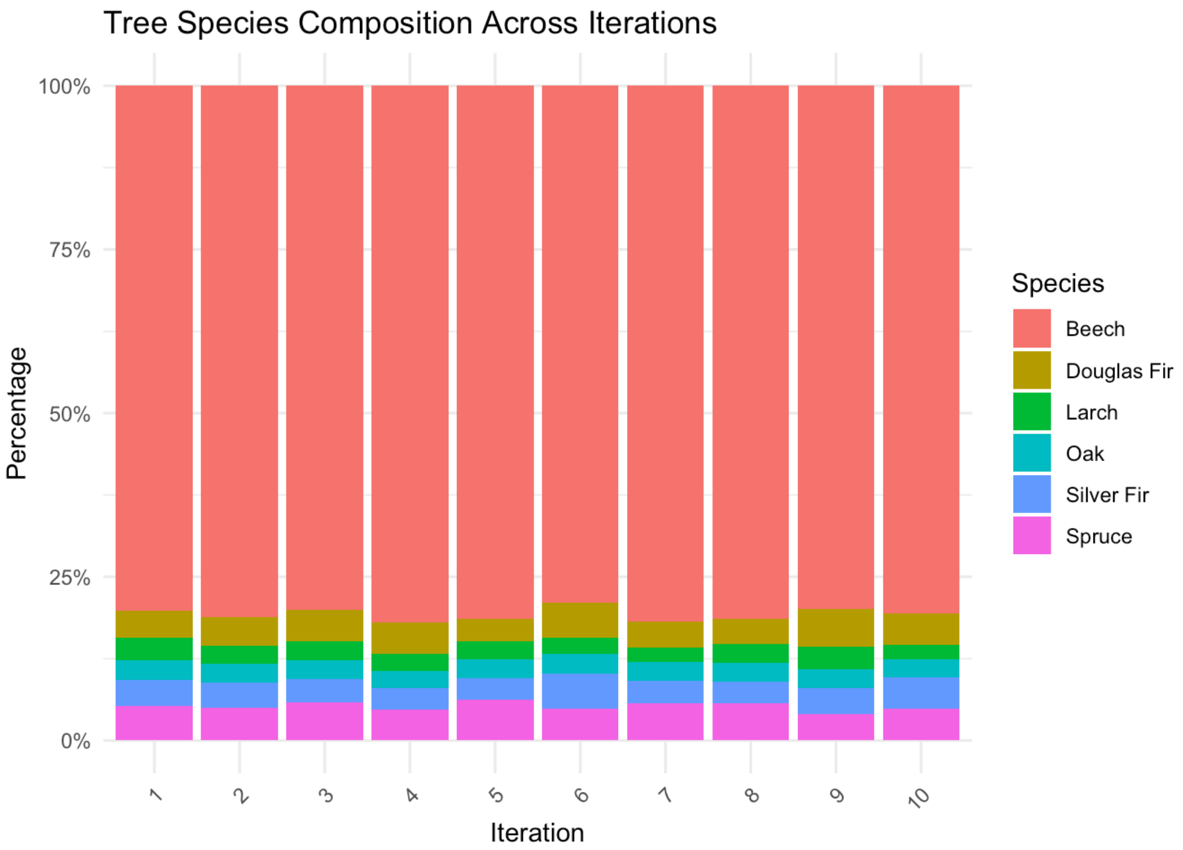
Viewing data in R “Real World Scenario“

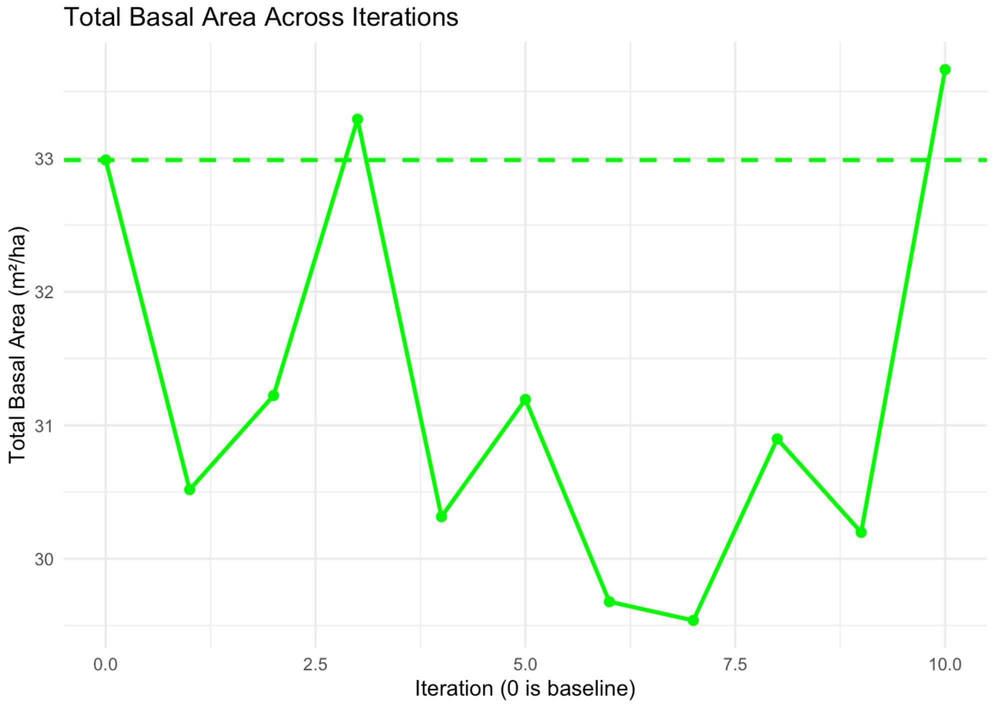
"Imagine you go into the forest to set points whose exact coordinates have already been determined. You navigate to the locations using your smartphone or a GNSS device. But in the forest the signal jumps and you are never quite sure whether you are standing at the exact point. Nevertheless, you start measuring because "position is position", right? Later in the office you notice that the points are slightly shifted, some by as much as 2 m. This may mean that you have missed some trees that were measured in previous years. "

We simulate!

(In reality, such measuring points are precisely staked out so that the exact plot coordinates can be found again. Nevertheless, we want to know whether this is actually that serious)

Viewing data in R Simulation of samples – GNSS uncertainties





→ Even a 2 m deviation from the actual plot has immense effects on the result

TAKE HOME – and what questions do you still have?

We now know
... what the Ecosense forest looks like (at least on paper)
... what sampling uncertainty is
... why this is important for forest inventories
... that there are random and systematic errors
... how we can quantify these errors
... how we can address the uncertainties
... how we can simulate forest inventories
... how we can represent geodata in R
... how we can distribute plots sensibly
... what the effects of (position-)inaccurate measurements are