× A linear allometric relationship is said to exist between two dimensional elements (X and Y) of an individual or population (in our case, a forest stand) when the relationship between them can be expressed in the following form:

$$Y = k X^a$$

- a allometric constant, characterizes the individual in a certain environment
- \checkmark k depends on the initial conditions and units of Y and X

- × The allometric equation therefore describes one tree variable as a function of other tree variable following the model.
- × Taking logarithms and then differentiating, it can be shown that the allometric equation assumes that the relative growth rate of one plant part is proportional to that of another

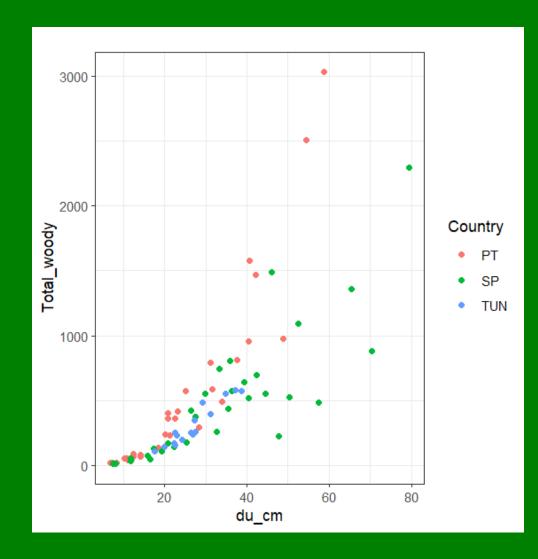
$$Y = k X^a \qquad \Leftrightarrow \qquad ln(Y) = ln(k) + a ln(X)$$

× The parameter *a* - allometric constant - is the coefficient of proportionality between the relative growth rates of the two plant parts

$$\frac{1}{Y}\frac{dY}{dt} = a\frac{1}{X}\frac{dX}{dt}$$

$$\int \frac{1}{Y} \frac{dY}{dt} = a \int \frac{1}{X} \frac{dX}{dt} \qquad \Rightarrow \qquad \ln(Y) = k + a \ln(X)$$

imes a can, therefore, provide direct information on the partitioning of assimilates between plant parts



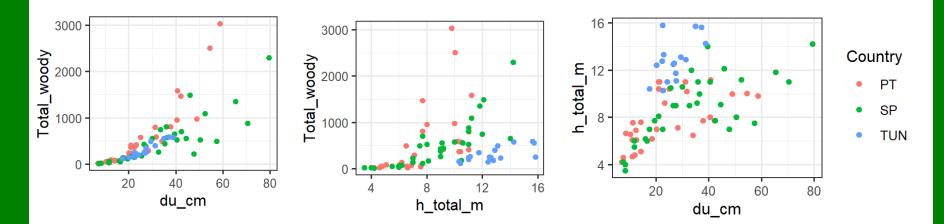
× The multiple allometric relationship is established between one tree variable and other tree variables (more than one):

 $Y = k X^a Z^b$

- × Allometric relationships can also be established between forest stand variables
- × The existence of allometric relationships between tree variables (or between stand variables) is very important for growth and yield modelling of trees and stands
- × It is one of the biologic hypothesis that can be used in the formulation of models

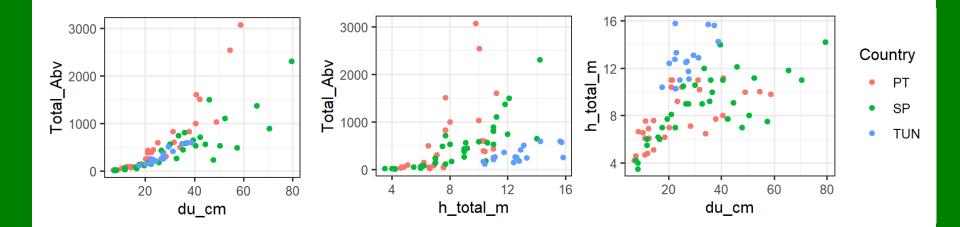
1. Initial data analysis

 Descriptive analysis of the data. Location and dispersion measures. Look for tendencies. Look for errors. Data limitations... ggplot, summary(dados), geom_boxplot()...



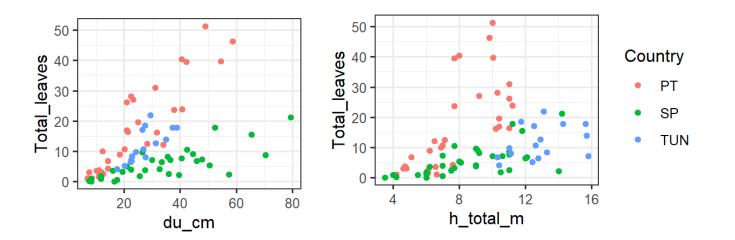
1. Initial data analysis

 Descriptive analysis of the data. Location and dispersion measures. Look for tendencies. Look for errors. Data limitations... ggplot, summary(dados), geom_boxplot()...

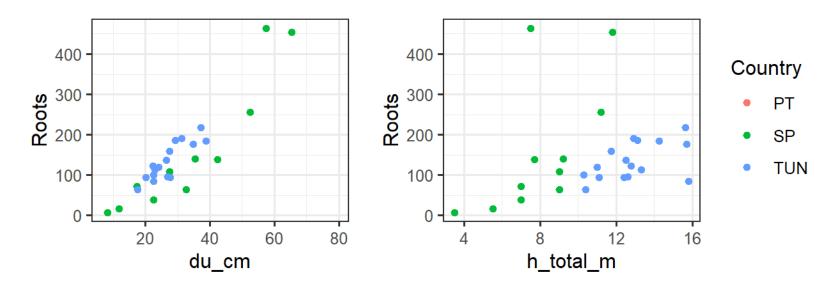


1. Initial data analysis

 Descriptive analysis of the data. Location and dispersion measures. Look for tendencies. Look for errors. Data limitations... ggplot, summary(dados), geom_boxplot()...



- 1. Initial data analysis
 - Descriptive analysis of the data. Location and dispersion measures. Look for tendencies. Look for errors. Data limitations... ggplot, summary(dados), geom_boxplot()...



2. Fit the allometric model. Total_woody = $a du^{b}$. Look at model parameter significance. Are they significantly different from zero in the model?

Statistical Tests for Parameter Significance: t-test for each parameter Bj (in this example a and b).

× The hypothesis for this test are:

 $H_0: \beta_j = 0$ (the parameter is not significant)

$$H_1: eta_j
eq 0$$

× The t-statistic is computed as:

 $t_j = rac{\hat{eta}_j}{\mathrm{SE}(\hat{eta}_j)}$ ~ t student distribution with (n-k) degrees of freedom under H₀.

2. Fit the allometric model. Total_woody = $a du^{b}$. Look at model parameter significance. Are they significantly different from zero in the model?

Statistical Tests for Parameter Significance: t-test for each parameter Bj (in this example a and b).

× The p-value is obtained as: $p = 2P(T > |t_j|)$

If p is small (e.g., p < 0.05), we reject H_0 , indicating that parameter eta_j is statistically significant.

2. Fit the allometric model: Total_woody = $a du^{b}$. Look at model parameter significance. Are they significantly different from zero in the model?

> summary(model) # checks the model

```
Formula: Total_woody ~ a * du_cm^(b)
```

```
Parameters:

Estimate Std. Error t value Pr(>|t|)

a 2.0095 1.2930 1.554 0.124

b 1.6013 0.1639 9.768 3.52e-15 ***

---

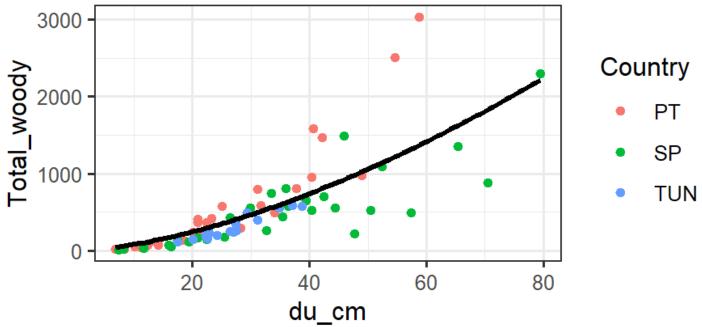
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 346.9 on 78 degrees of freedom

Number of iterations to convergence: 9

Achieved convergence tolerance: 1.745e-06
```

3. Look at model parameter estimates. Interpret them under biological meaning. Plot the estimates on the data.

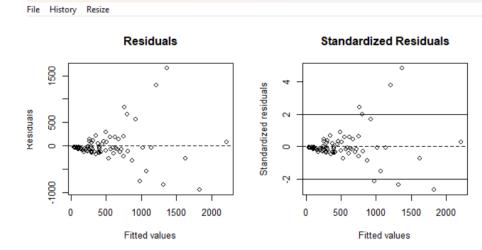


R Graphics: Device 4 (ACTIVE)

4. Verify model conditions

- Residuals distribution
- # residuals plots

plot(nlsResiduals(model))



Х

4. Verify model conditions

- Normality of the residuals
- # normality test

shapiro.test(dados\$Total_woody)

The hypotheses for the Shapiro-Wilk test using normal distribution notation are:

• Null hypothesis (H_0): The data follows a normal distribution.

 $H_0: X \sim N(\mu, \sigma^2)$

• Alternative hypothesis (H_1): The data does not follow a normal distribution.

$$H_1: X
ot \sim N(\mu, \sigma^2)$$

If the **p-value** is small (typically p < 0.05), we reject H_0 , suggesting that the data does not follow a normal distribution. If $p \ge 0.05$, we fail to reject H_0 , meaning there is no strong evidence against normality.

- 4. Verify model conditions
 - Normality of the residuals
 - # normality test

```
shapiro.test(dados$Total_woody)
```

> shapiro.test(dados\$Total_woody) # normality test

Shapiro-Wilk normality test

data: dados\$Total_woody
W = 0.7272, p-value = 6.833e-11

4. Verify model conditions

✓ Normality of the errors

QQ plot (Quantile-Quantile plot) - graphical tool used to assess whether a dataset follows a specific theoretical distribution, such as a normal distribution. It compares the quantiles of the observed data with the quantiles of the theoretical distribution.

How it works:

- Quantiles of the observed data are plotted on the y-axis.
- Quantiles of the theoretical distribution (e.g., normal distribution) are plotted on the x-axis.
- The points on the plot represent pairs of quantiles from both datasets (observed vs. theoretical).

4. Verify model conditions

Normality of the errors

QQ plot (Quantile-Quantile plot) - graphical tool used to assess whether a dataset follows a specific theoretical distribution, such as a normal distribution. It compares the quantiles of the observed data with the quantiles of the theoretical distribution.

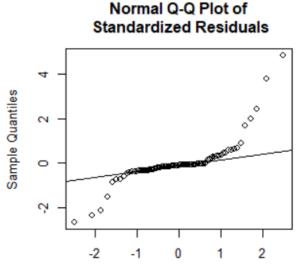
If the data is normally distributed (for example), the points will lie approximately on a straight line (usually a 45-degree line or a line with slope 1). If the points deviate significantly from the straight line, it suggests that the data does not follow the chosen theoretical distribution.

4. Verify model conditions

✓ Normality of the errors

How to plot the QQ plot

Option 1) plot(nlsResiduals(model)))



Theoretical Quantiles

4. Verify model conditions

✓ Normality of the errors

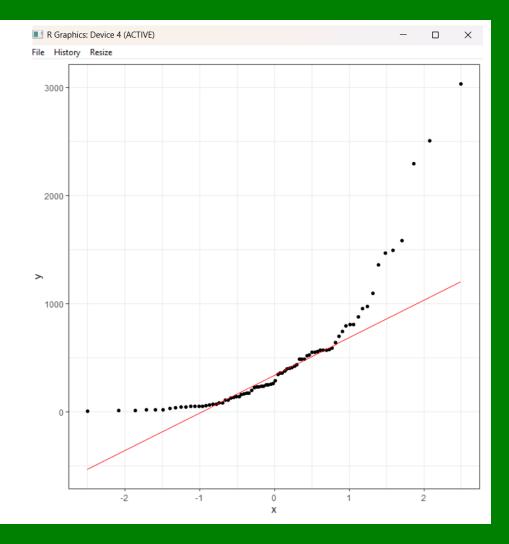
How to plot the QQ plot

Option 2) ggplot

```
# check normality of y
ggplot(dados, aes(sample = Total_woody)) +
   stat_qq() + stat_qq_line(colour = "red") +
   theme_bw()
```

4. Verify model conditions

Normality of the errors
 How to plot the QQ plot
 Option 2) ggplot



4. Verify model conditions

Normality of the errors

Key interpretations of a QQ plot:

- Straight line: The data follows the specified distribution (e.g., normal).
- Curved pattern: Indicates skewness in the data.
- If the points bend upward at the ends, it suggests that the data has heavy tails (more extreme values than the normal distribution).
- If the points bend downward at the ends, it suggests light tails (fewer extreme values than the normal distribution).
- S-shaped or other patterns: Indicates that the data may not match the specified distribution at all.

4. Verify model conditions

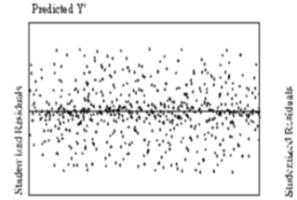
Homoscedasticity: denotes the assumption that the variance of the residuals remains constant across all levels/range of the predicted/estimated values of the dependent variable. Put simply, it signifies that the dispersion of residuals stays consistent, enhancing the accuracy and legitimacy of regression predictions. Heteroscedasticity, the violation of homoscedasticity, is present when the residual values varies across values of predicted/estimated values of the dependent variable.

- 4. Verify model conditions
 - Homoscedasticity: residuals (yy) versus predicted values (xx)

Homoscedasticity

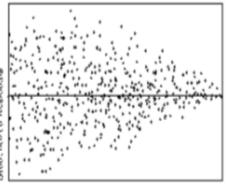
Heteroscedasticity

Heterosced asticity

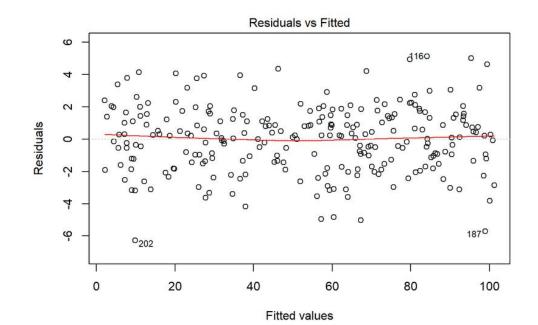






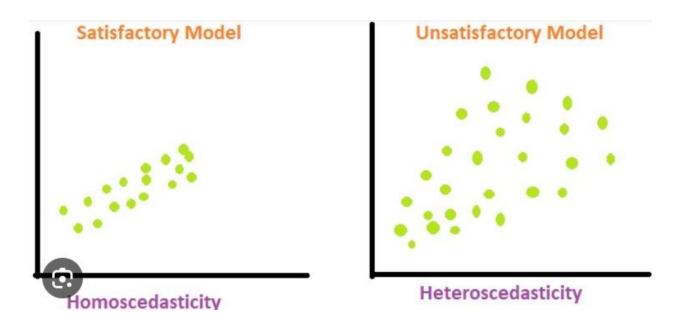


- 4. Verify model conditions
 - Homoscedasticity: residuals (yy) versus predicted values (xx)



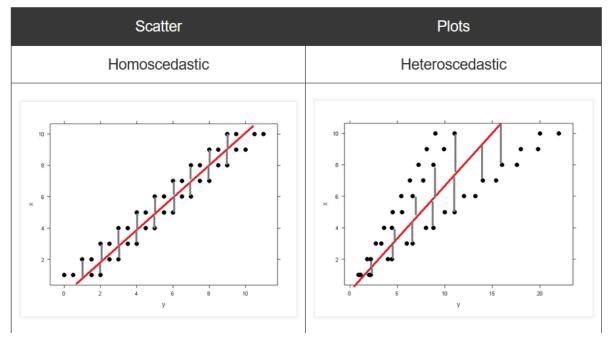
4. Verify model conditions

 Homoscedasticity: Observed values (yy) versus predicted values using the fitted model (xx)



4. Verify model conditions

 Homoscedasticity: Observed values (yy) versus predicted values using the fitted model (xx)



Additional reading

- × MYERS, R. H. 1990. Classical and Modern Regression with Applications. Second edition. Duxbury Classic Series. Chapters 3, 5 and 6.
- X Jorge, C.; Tomé, M.; Ruiz-Peinado, R.; Zribi, L.; Paulo, J.A. Quercus suber Allometry in the West Mediterranean Basin. Forests 2023, 14, 649. <u>https://doi.org/10.3390/f14030649</u>
- X A.C. Correia, S.P. Faias, R. Ruiz-Peinado, F. Chianucci, A. Cutini, L. Fontes, M.C. Manetti, G. Montero, P. Soares, M. Tomé. 2018. Generalized biomass equations for Stone pine (Pinus pinea L.) across the Mediterranean basin. Forest Ecology and Management, Volume 429, Pages 425-436, <u>https://doi.org/10.1016/j.foreco.2018.07.037</u>
- https://ademos.people.uic.edu/Chapter12.html