



Universidad de Valladolid

Master Erasmus Mundus in Mediterranean Forestry and
Natural Resources Management
(MEDfOR)

Tree biomass allocation in temperate mixed forests

Eric Cudjoe

Dr. Felipe Bravo Oviedo

Supervisor

Dr. Ricardo Ruiz-Peinado

Co-supervisor

July 2019



Co-funded by the
Erasmus+ Programme
of the European Union



Escuela Técnica Superior
de Ingenierías Agrarias **Palencia**



INDEX

| | |
|--|-------------------------------------|
| LIST OF FIGURES..... | 5 |
| LIST OF TABLES..... | 6 |
| LIST OF ANNEXES..... | Error! Bookmark not defined. |
| LIST OF ABBREVIATIONS | 7 |
| RESUMEN | 8 |
| ABSTRACT..... | 9 |
| 2. OBJECTIVES..... | 13 |
| 3. MATERIALS AND METHODS | 14 |
| 3.1. Study area | 14 |
| 3.2. Measurements and data collection..... | 16 |
| 3.3. Data analysis | 18 |
| 3.3.1. Allometry analysis | 18 |
| 3.3.1.1. Height to diameter equation | 18 |
| 3.3.2. Biomass equation..... | 20 |
| 3.3.2.1. Selection of predictors..... | 20 |
| 3.3.2.2. Biomass allometry of monospecific and mixed stands | 20 |
| 3.3.2.3. Estimation of total aboveground biomass..... | 21 |
| 3.3.2.4. Estimation of component biomass | 21 |
| 3.3.2.5. Evaluation | 22 |
| 4. RESULTS..... | 22 |
| 4.1. Stand characteristics..... | 22 |
| 4.1.1. Height and diameter a breast height equation..... | 22 |
| 4.2. Biomass analysis | 23 |
| 4.2.1. Variables selection for Aboveground Biomass Estimations | 23 |
| 4.2.2. Relationship between independent variables for monospecific and mixed stands..... | 23 |
| 4.2.3. Differences in biomass allometry of monospecific and mixed stands | 24 |
| 4.2.4. Estimation of biomass component proportions using Dirichlet method | 25 |
| 4.2.5. Estimation of total aboveground biomass..... | 26 |
| 5. DISCUSSION | 28 |
| 6. CONCLUSION..... | 31 |

7. ACKNOWLEDGEMENT32

8. REFERENCES33

ANNEX37

1. Supplementary figures **Error! Bookmark not defined.**

2. Supplementary tables **Error! Bookmark not defined.**

3. R scripts37

LIST OF FIGURES

- Figure 1. Map of the study area in the North of Palencia and the North of Burgos.....**Error! Bookmark not defined.**
- Figure 2. Classification of aboveground component biomass Data and fitted Korf curve model (model no.3) for the different plots in the triplets**Error! Bookmark not defined.**
- Figure 3. Data and fitted Korf curve model (model no.3) for the different plots in the triplets **Error! Bookmark not defined.**
- Figure 4. The relationships among independent variables for monospecific and mixed stand trees **Error! Bookmark not defined.**
- Figure 5. Differences in allometry of monospecific and mixed stand trees**Error! Bookmark not defined.**
- Figure 6. The trend in component biomass proportion with respect with to DBH..**Error! Bookmark not defined.**

LIST OF TABLES

| | |
|--|----|
| Table 1: General stand characteristics for the two studied sites (Palencia and Burgos, Northern Spain)..... | 15 |
| Table 2: Summary statistics of the harvested tree data used in Burgos study area..... | 18 |
| Table 3: Height-diameter models selected for performance test..... | 19 |
| Table 4: Biomass models tested for total aboveground biomass for both species..... | 21 |
| Table 5: Parameter estimates of estimate for equation [3] in Table 3..... | 23 |
| Table 6: Parameter estimates and their standard errors for the In-likelihood model to estimate total aboveground biomass in Scots pine and sessile oak trees..... | 26 |
| Table 7: Parameter estimates and the standard errors for the component biomass models using the Dirichlet regression where Scots pine includes DBH (cm) tree (Ht, m) and sessile oak includes only Ht (m)..... | 27 |
| Table 8: Evaluation statistics root mean square error (RMSE), RMSE percent, bias and bias percent in component biomass estimation in Scot pine and sessile oak trees using Dirichlet regression..... | 28 |

LIST OF ABBREVIATIONS

DBH: Diameter at breast height at tree level

AGB: Above-ground biomass

ATB: Aboveground total biomass

Ht: Total height

TKB: Thick branches

TNB: Thin branches

ST: Stem

B1: Thin branches

B2: Thick branches

CR: Crown ratio

UL: Crown length

CBH: Crown base height

ANCOVA: Analysis of Covariance

Bravo 1: 1st triplet in Palencia site

Bravo 2: 2nd triplet in Palencia site

Rio 1: 1st triplet in Burgos site

Rio 2: 2nd triplet in Burgos site

RESUMEN

En este siglo XXI, uno de los principales aspectos en la gestión sostenible de los recursos son las estimaciones de las reservas forestales de carbono y los balances de carbono para las diferentes especies tanto a nivel global como local. Este trabajo de investigación se centra principalmente en ilustrar cómo la mezcla de especies en los rodales forestales afecta las relaciones alométricas de biomasa y la distribución de la misma en rodales monoespecíficos y mixtos en *Pinus sylvestris* L. y *Quercus petraea* (Matts.) Liebl. En este estudio, ajustamos 4 modelos diferentes de altura y diámetro para bosque monoespecífico y mixto para especies de pino silvestre y roble albar. La curva de Korf para los modelos de altura y diámetro se comportó mejor que otros modelos como los valores de AIC más bajo y R^2 más alto sugieren. Además, no se encontró diferencias en la relación altura diámetro en los robles entre las masas puras y mixtas. Si embargo, se encontraron diferencias en la relación altura-diámetro entre masas puras y mixtas en el caso del pino silvestre. El desarrollo de modelos de altura-diámetro que estiman la altura del árbol utilizando un subconjunto de árboles muestreados presenta un enfoque para complementar los muestreos donde solo se ha medido el diámetro.

El análisis de ANCOVA demostró que la alometría de la biomasa de árboles del monoespecífico muestra un contraste con el soporte mixto, pero la diferencia no es significativa.

Se usó la regresión de ajuste Dirichlet para seleccionar los mejores modelos de biomasa con el diámetro a la altura del pecho (DBH) y la altura total (Ht) como variables / predictores independientes para ambas especies. La biomasa aérea total de rodales mixtos no fue significativamente diferente de los rodales monoespecíficos. Esto implica que la proporción de biomasa en diferentes componentes de árboles de rodales mixtos es similar a la de árboles de rodales monoespecíficos. Los resultados muestran que la mayor proporción de biomasa, para las 2 especies analizadas, se encuentra en el fuste y la menor en el follaje. Por lo tanto, los modelos alométricos de biomasa desarrollados a partir de rodales monoespecíficos se pueden utilizar para predecir la biomasa de árboles en rodales mixtos sin un error importante.

Palabras clave: pino silvestre, roble albar, regresiones Dirichlet, biomasa aérea, distribución de biomasa,

ABSTRACT

At this 21st century, one of the key landmarks in sustainable resource management is the estimations of forest carbon stocks and carbon balances at different geographical ranges and coming up with efficient forestry solutions to mitigate climate change. Thus, the comprehension of forest carbon cycles and dynamics for sustainable resource management needs quantification of forest biomass at the species both global and local level. This research work was mainly focused on illustrating how the forest stand mixture affect biomass allometric relationships and allocation of biomass in monospecific and mixed stands of *Pinus Sylvestris* L. and *Quercus petraea* (Matts.) Lieb. in Northern Spain. In this study, we fit 4 different height-diameter models for monospecific and mixed forest for Scots pine and Sessile oak species. The Korf curve for height-diameter models performed better than other models as the lowest AIC and highest R² obtained suggest. In addition, there is a no variation of height-diameter relationships for oak trees in monospecific or mixed stands. However, an extensive difference in the monospecific and mixed stands for Scots pine was found. The development of height-diameter models that estimate tree height using a subset of sampled trees presents an approach to supplement surveys where only diameter has been measured.

ANCOVA analysis proved that the tree biomass allometry of the monospecific show contrast from the mixed stand but the difference is not significative. Dirichlet fitting regression was used to select the best biomass models with the diameter at breast height (DBH) and total height (Ht) as independent variables/predictors for both species. the total aboveground biomass of mixed stands was not significantly different from monospecific stands. Also, the proportion of biomass in different tree component of mixed stand trees is similar to that of monospecific stand trees. The results illustrate that the largest biomass allocation for the 2 species was on stem and the least was on foliage. Therefore, the biomass allometric models developed from monospecific stands can be used to predict tree biomass in mixed stands without an important error.

Keywords: Scots pine, Sessile oak, Dirichlet regressions, aboveground biomass, biomass allocation

1. INTRODUCTION

The forests allocate great amounts of carbon in their plant tissues, playing a key role in controlling the atmospheric CO₂ balance on earth (Wang, Fang, & Zhu B, 2008), thus there is the need to estimate its biomass, and the relationships among covariables. As there is an increased interest in carbon accounting in forest ecosystems, there is a need for efficient methods to estimate the above- as well as the below-ground biomass components of the trees (Næsset, 2004). However, because of the high costs and time consuming efforts to estimate below-ground biomass, nearly all carbon-associated research is limited to above-ground biomass (Yuen et al., 2013; Ziegler et al., 2012). For this reason, majority of the research works regarding tree biomass use allometric equations which has been developed out of the traditional harvesting and weighing of trees. Usually the allometry equations are built on aboveground biomass because the forest ecosystems capture most of the carbon pools in the aboveground biomass (Vashum & Jayakumar, 2012). In addition, Czapowskyj et al., (1985) and Xiao et al., (2003) found that 80% or more of biomass contain in the above-ground components compared to belowground. This explains that the above-ground biomass is the central point of carbon storage. As a result, most studies focus on above-ground neglecting the belowground biomass that necessitates uprooting of tree roots but there may be some differences between conifers and broadleaves species.

Scots pine (*Pinus sylvestris* L.) being the most widely distributed of the pine species and amongst the most important timber species in Eurasia (Stanners & Bourdeau, 1995; Mason & Alía, 2000) due to its ability to thrive on different substrates, well-developed or poor soils, in cool wet climates, or in very cold to extremely cold continental climates (Cañellas et al., 2000). Stanners & Bourdeau (1995) stated that Scots pine forests is an important part of Europe's carbon budget covering 24% of the total forested area (75 million km²). Also, *Quercus petraea* (Matt.) Liebl., (sessile oak) are common broadleaved tree species in Europe.

Scots pine and sessile oak are two main forest species in the temperate regions not only because of their widest distribution but also their vital socio-economic and ecological importance. Since both species grow well with other species, it is essential to know the effect of mixture for such composition. Recently, a number of research has been conducted on mixed stands (Dutca et al., 2017) but none of the research work has focused on mixed stands of Scots pine and Sessile oak. Mixture could be more advantageous than monospecific stands depending on the composition of mixture. Research by (Pretzsch, 2018) affirmed that forest dynamics including growth and yield is positively influenced by forest composition (monospecific or mixed stands). Furthermore, several studies have supported and shown evidence of the advantages of mixed forest to monospecific

forests (Gamfeldt et al., 2013; Pretzsch et al., 2015). Therefore, this study could help in knowing whether Scots pine admixture with Sessile oak exhibit aboveground competition, reduced competition or facilitation especially in terms of light availability. In terms of biomass production, much of the biomass research has been conducted in monospecific forests, being most of the studies limited to local or regional sites and not a complete set of forest species were studied.

Most of the forest managers have even moved their focus towards mixed forest presently (Lu et al., 2018) as a way of adaptation against the climate change. Thus, the practice of slowly moving from monospecific forests to mixed forests has now become very popular in most areas in Europe (Agnoletti, 2006) as a result of the positive benefits provided by mixed forest in productivity, stability, and ecosystem services (Keenan, 2012; Bauhus. et al., 2013; del Río et al., 2017). Mixed forests had received immensely attention in forestry due to its crucial influence in species diversity, forest adaptation and mitigation of climate change.

The concept of monospecific forests is well studied and unveil whereas scientific advances about mixed forests are recent, some knowledge gaps have been identified (Coll et al., 2018). For example, the use of existing biomass models developed mainly in monocultures forests. These models should be tested and checked for accuracy in mixed forests. However, if monospecific models are not suitable to mixed forests new biomass species-specific models should be developed and fitted for mixed stands. The outcome of the few studies performed on mixed forests is difficult to generalize because of the changes in the forest structure (Forrester & Pretzsch, 2015). Most of the studies affirm that, mostly encouraging mixed-species forests can be an appropriate and convenient alternative to cope with climate change-induced impacts. Although there are other studies regard to mixed stands, example del Rio & Sterba, (2009) found the contrary.

The changes of height-diameter relationships mostly depend on the species composition and the environmental conditions of the region (Augusto da Silva Scaranello et al., 2012). Tree height is usually used as an independent variable for forest models, but due to its high-cost and difficulty in measuring the height. Site-specific height-diameter models is used as the best option for height measurements (Batista, Couto, & Marquesini, 2001).

Most of the studies on mixing effects pay attention primarily on stand density and species composition (del Río et al., 2018). Therefore, more research work needs to show stand dynamics in mixed forests including mixed-effect models in the prediction of height-diameter (H-D) relationship. It is obvious that height to diameter functions are so important in predicting height growth (Augusto

da Silva Scaranello et al., 2012), and also in many growth and yield models (Augusto da Silva Scaranello et al., 2012); Huang et al., 1992). Therefore, they need not to be neglected during studies. So, one of the main questions is whether the mixed forests have different tree biomass allocation and tree allometry to monospecific forests in order to improve our knowledge regarding stand dynamics of tree mixture. Nevertheless, we need to take into account the species composition and species proportion of mixture (Légaré, Paré, & Bergeron, 2004) in order that our results can be generalized. Many research works have been done comparing monospecific and mixed stands, but it is still not well known how tree biomass models are influenced by forest composition. It is essential to understand and address the changes in tree biomass allometry due to tree mixture and its implication on silviculture and carbon sequestration in order to mitigate climate change (Dutca et al., 2017). Moreover, complementarity effects between species mixtures could cause changes in allometry due to distinct utility of available resources due to the differences in temperament (shade tolerance) or rooting pattern (shallower or deeper).

Using forest inventory data for carbon estimation often depends on species-specific biomass equations and/or biomass expansion factors which is obtained with the help of empirical data. Hence, models development need sample tree harvesting, measurement and also determination of the dry weight of various biomass components (Correia et al., 2018). There are numerous studies that developed allometric biomass equations for several forest tree species all over the world (Pastor et al., 1984; Zianis et al., 2005; Henry et al., 2011; Mosseler et al., 2014; Luo et al., 2019). The most prevailing tree variables used in biomass model fitting is diameter at the breast height, 1.30 m from stem base (Karlik & Chojnacky, 2014) as it is usually measured in all forest inventories. Other variables that have been used in biomass models are including total tree height, crown length, basal diameter, or wood specific gravity to include species identity (Hunter et al., 2013; Burquez et al., 2014; Dahlhausen et al., 2017). Forest modelling was improved with the introduction of the use of seemingly unrelated regression, SUR (Zellner, 1962; Zellner & Theil, 1962). The generalization of this technique for fitting different tree component biomass equations mainly stem, branches and foliage (Parresol, 1999); (Parresol, 2001) allows to provide accurate biomass models. This is because it ensures additivity between components and total biomass predictions thus, assure that the total tree aboveground biomass is the sum total of the aboveground component estimations (Balboa-Murias et al., 2006; Ruiz-Peinado et al., 2011, 2012; Tesfaye et al., 2016; Correia et al., 2018). Nevertheless, other methodologies for fitting are based in predicting the proportions of biomass in each component using Beta regression, Dirichlet regressions or Multinomial In-linear regression (Poudel & Temesgen, 2016a); Eker et al., 2017). In order to obtain

the amount of biomass for every component, the predicted proportions are applied to the observed total aboveground biomass (Poudel & Temesgen, 2015) . The Dirichlet regression, which has been recently tested in order to observe the additivity property between biomass compartments and total aboveground biomass, will be used in this study as in some works has been identified as better than SUR fitting and it is implemented in R 3.5.3 software by using package 'DirichReg' (Maier, 2014).

We may be confronted by some crucial limitations even though the functions used are relatively simple but could present some issues. Firstly, the functions are not particularly generic: if we change species, the equations must be used with caution. Also, other limitations about the biomass study can be related with the small number of harvested trees that could influence the generalization of the results, although the samples are collected along the diameter range found in the stands.

2. OBJECTIVES

The essential objective of this study is to know if the stand dynamics of trees growing in Scots pine (*Pinus sylvestris* L.) monocultures, sessile oak (*Quercus petraea* (L.) H. Karst) monospecific stands and in Scots pine - sessile oak mixed stands differ. For this objective, some allometric relationships and the biomass allocation pattern of the species will be studied in both monospecific and mixed stands.

However, to make the studies feasible, certain number of specified objectives need to be assigned to this work:

- ✓ To examine the relation of DBH-H equation and biomass allometry in monospecific and mixed stands.
- ✓ To predict the total aboveground biomass and proportion of biomass in each component to account for the differences and changes among the stands.
- ✓ To develop tree biomass models for monospecific Scots pine (*Pinus sylvestris*.) stands, monospecific sessile oak (*Quercus petraea*) stands and Scots pine - sessile oak mixed stands for different biomass components (stems, thick and thin branches, and foliage) and to determine whether the estimated models differ from its monospecific and mixed stands.

3. MATERIALS AND METHODS

3.1. Study area

This study was performed in two forest stands, located at the north of Burgos (Busnela municipality) and north of Palencia (municipality of Aguilar de Campoo) in the Cantabrian Mountains. In these forest stands, two thinning trials were installed by the INIA-CIFOR (Burgos site) and the University of Valladolid (Palencia site). Every site includes also a control (no thinning) and a thinning experiment. The thinning sites were implemented in research collaboration between European institutions under the umbrella of the REFORM-mixing project (<http://www.reform-mixing.eu/>). Both sites were designed with the methodology called triplet that has been used by this publication (Heym et al., 2017). Each triplet contained a monospecific Scots pine plot, a monospecific sessile oak plot and a mixed plot. Every plot shows similar site characteristics (soil conditions, site quality, environmental and silvicultural conditions etc.) and similar DBH range. Fortunately, the triplet experiment provides more information on the mixing effect of species precisely Scots pine and sessile oak. The stand in Burgos was established as a plantation of Scots pine and natural regeneration of sessile oak, while the stand in Palencia consisted of natural regeneration of both species. Both locations have trees of the same ages in its pure and mixed stands. The elevation of both stand sites ranges from approximately 760-1300 m and with sandy loam as soil type. It has a cold and temperate climate. One peculiar case was the nomenclature of the triplet, Bravo 1 was the first triplet in Palencia and Bravo 2 was the second triplet in Palencia. On the other site, Rio 1 was the first triplet in Burgos and Rio 2 was the second triplet in Burgos.

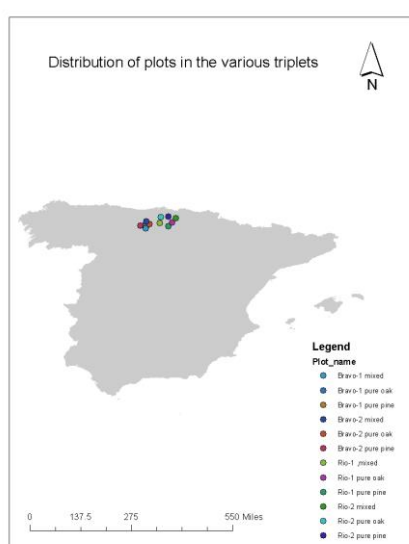


Figure 1: Map of the study area in the North of Palencia and the North of Burgos

Table 9: General stand characteristics for the two studied sites (Palencia and Burgos, Northern Spain)

| Site | Triplet | Plot | Long (°W) | Lat (°N) | Alt (m) | Slope (%) | Mean DBH | Mean Ht | Area (ha) | N | BA |
|-----------------|----------------|-----------|---------------|----------------|---------|-----------|----------|---------|-----------|------|-------|
| Palencia | Bravo 1 | Pure pine | 04°14'31.98" | 42°53'52.09" | 1066 | 16 | 26.92 | 17.43 | 0.062 | 1134 | 71.30 |
| | | Pure oak | 04°14'22.60" | 42°53'43.47" | 1512 | 10 | 17.89 | 15.29 | 0.062 | 1240 | 58.70 |
| | | Mixed | 04°14'32.22" | 42°53'41.91" | 1188 | 11 | 23.51 | 17.69 | 0.089 | 1165 | 54.20 |
| | Bravo 2 | Pure pine | 04°14'36.02" | 42°53'51.04" | 1185 | 17 | 29.91 | 18.05 | 0.091 | 869 | 66.40 |
| | | Pure oak | 04°14'30.34" | 42°53'40.25" | 1188 | 15 | 19.71 | 17.59 | 0.060 | 1468 | 44.00 |
| | | Mixed | 04°14'31.66" | 42°53'31.48" | 1188 | 15 | 23.82 | 18.05 | 0.086 | 1238 | 55.10 |
| Burgos | Rio 1 | Pure pine | 03° 47'19.19" | 43° 2' 55.17 " | 810 | 22 | 28.01 | 15.27 | 0.062 | 1102 | 66.17 |
| | | Pure oak | 03° 47'21.36" | 43° 2' 51.61" | 760 | 34 | 18.36 | 14.6 | 0.058 | 1461 | 47.65 |
| | | Mixed | 03° 47'19.26" | 43° 2' 52.64" | 785 | 28 | 23.47 | 16.09 | 0.089 | 1203 | 52.65 |
| | Rio 2 | Pure pine | 03° 47'17.64" | 43° 2' 54.10 " | 815 | 20 | 27.58 | 14.11 | 0.090 | 847 | 64.00 |
| | | Pure oak | 03° 47'24.34" | 43° 2' 58.39 " | 765 | 35 | 21.53 | 15.62 | 0.058 | 1433 | 41.96 |
| | | Mixed | 03° 47'23.23" | 43° 2' 55.84 " | 780 | 24 | 22.85 | 16.18 | 0.086 | 1203 | 54.95 |

N, number of stems per hectare including all trees within the said plot; Long, longitude; Lat, latitude; DBH, diameter at breast height in centimeters; H, height in meters; Alt, altitude; Bravo and Rio are the names triplet in the different sites.

3.2. Measurements and data collection

Date of survey for the stands in Palencia and Burgos were 20th September 2017 and 12th April 2018 respectively. First of all, the stand variables including the DBH (cm), total height (m), height of the crown base(m), crown width (m) and tree position were measured Secondly, thinning were carried out only in the Burgos site during the spring of 2018 and during the thinning operations biomass data was collected for some trees. Unfortunately, thinning in the Palencia site was not carried out due to problems between the Forest Service and the forest enterprise; therefore, above-ground biomass estimation was performed only on Burgos site. In a nutshell, the work was divided into two parts. The first part comprises of using the data from both sites to develop height-diameter allometry equations for monospecific and mixed stands. The second part consists of using the available biomass data from Burgos site to develop above-ground estimations. Firstly, we sampled all the DBH of the trees and measured tree heights within the 4 triplets containing 12 plots. Only height and diameter measurements are needed to build and compare the allometry relationship for both species. Secondly, at the Burgos site, 8 trees per species per plot were selected in the monospecific and mixed stands. Thus, the total number of trees sampled was 32 trees. The trees were selected upon the criteria listed: (a) trees were selected along the diameter range found in each stand; (b) suppressed trees or trees that shows signs of affected disease or structural integrity were refrained in all the stands; (c) in the mixed forest, sampled trees must be in different species competition conditions in order to test the inter-specific competition settings). Trees were cut and different biomass components were accounted: stem with bark, thick branches, thin branches and foliage (Montero, Ruiz-Peinado, & Muñoz, 2005); (Ruiz-Peinado, del Rio, & Montero, 2011),(Ruiz-Peinado et al., 2012). Crown biomass components were totally stripped from the trunk after chopping of each tree sample. The foliage biomass was stripped from the branches according to the sizes of the tree sample. Crown length (CL) was calculated as the difference between total tree height (Ht; m) and height to the base of the crown (HBC; m), crown radii (CR; m) was also calculated based on the Crown length (CL) divided by total tree height (Ht; m). The fresh biomass of each component was measured in the field using an electronic scale and a representative subset sample of each of them was collected and in laboratory was dried at 102 °C to constant mass. The dried biomass portion of other components (stem, thick branches, thin branches and foliage) was finally calculated. In order to obtain the biomass for each tree sample compartment (including stem, thick branches, thin branches and foliage).

The moisture content was calculated: Total dry weight (kg) = Total fresh weight × (Sample dry weight/ Sample fresh weight). As it is not possible to weigh the stem, its diameter was recorded at 1m intervals to obtain the volumes using Smalian's formula. Wood density was used to calculate stem dry biomass weight (Gutiérrez & Plaza, 1967). Dry biomass data was used for comparing species growing in monospecific stands and mixed stands.



Foliage: needles or leaves;

Thin branches: small part of the branches with a diameter lower than 7 cm;

Thick branches: diameter greater than 7 cm;

Stem: the huge part of the tree which is the commercial volume to a top diameter of 7cm.

Figure 2: Classification of aboveground component biomass

Furthermore, the summary statistics of the harvested tree data in Burgos study area for the estimation of total aboveground biomass and its components are given in Table 2.

Table 10: Summary statistics of the harvested tree data used in Burgos study area

| Pine | | | | | | | | |
|-------------------|---------|--------|---------|--------|---------|--------|---------|--------|
| | Pure | | | | Mixed | | | |
| Variable | Minimum | Mean | Maximum | SD | Minimum | Mean | Maximum | SD |
| Stem (kg) | 71.05 | 175.27 | 304.12 | 80.54 | 72.30 | 176.30 | 374.90 | 97.38 |
| Foliage (kg) | 3.13 | 6.55 | 11.10 | 2.67 | 0.70 | 4.01 | 11.58 | 3.57 |
| Thin Branch (kg) | 7.54 | 15.78 | 26.75 | 6.42 | 1.33 | 7.62 | 22.00 | 6.79 |
| Thick Branch (kg) | 3.41 | 23.78 | 50.48 | 18.24 | 4.52 | 19.22 | 39.70 | 13.34 |
| Total (kg) | 85.37 | 221.39 | 384.59 | 105.71 | 78.85 | 207.12 | 448.17 | 117.49 |
| DBH (cm) | 17.25 | 24.65 | 31.80 | 5.30 | 17.15 | 24.19 | 33.30 | 4.94 |
| Total height (m) | 14.9 | 15.86 | 17.00 | 0.79 | 14.95 | 17.62 | 21.60 | 2.30 |
| Oak | | | | | | | | |
| | Pure | | | | Mixed | | | |
| Variable | Minimum | Mean | Maximum | SD | Minimum | Mean | Maximum | SD |
| Stem (kg) | 15.42 | 110.06 | 277.44 | 84.66 | 14.06 | 92.31 | 222.27 | 71.09 |
| Foliage (kg) | 0.23 | 1.41 | 3.84 | 1.31 | 0.32 | 1.17 | 2.69 | 0.92 |
| Thin Branch (kg) | 0.49 | 3.01 | 2218.00 | 2.80 | 0.81 | 2.97 | 6.84 | 2.35 |
| Thick Branch (kg) | 5.05 | 14.74 | 36.76 | 10.90 | 7.78 | 20.76 | 43.86 | 13.49 |
| Total (kg) | 24.10 | 129.22 | 326.26 | 98.97 | 22.96 | 117.21 | 275.66 | 86.33 |
| DBH (cm) | 8.05 | 14.14 | 22.00 | 4.60 | 8.30 | 13.83 | 20.85 | 4.18 |
| Total height (m) | 11.30 | 15.38 | 18.60 | 2.46 | 11.40 | 15.60 | 18.55 | 2.06 |

DBH: diameter at breast height (1.3m); SD: standard deviation

3.3. Data analysis

3.3.1. Allometry analysis

3.3.1.1. Height to diameter equation

There are several studies that explain height and diameter relationships for different species and areas but limited information on height-diameter models (Batista et al., 2001). However, in this study, we fit height-diameter models for monospecific and mixed forests in Scots pine and sessile oak. We tested the performance of 4 (four) height-diameter models (Table 3) and we were able to select the best model using the standard error of residuals of the model. However, the overall best model(s) was fitted for the best site-specific height diameter model using the functions “Nonlinear Least Squares”. Then we can compare the pooled models selected for monospecific and mixed stands for the 2 species.

Table 11: Height-diameter models selected for performance test

| No | Reference | Model |
|----|-------------------|--------------------------------|
| 1 | Power equation | $H = a \cdot D^b$ |
| 2 | Meyer equation | $H = a \cdot (1 - \exp^{-bD})$ |
| 3 | Korf curve | $H = a \cdot \exp^{-bD^c}$ |
| 4 | Naslunds equation | $H = 1.3 + D^2 / (a + bD)^2$ |

Note: H = total tree height (m); D = DBH (cm); a, b and c are parameters to be estimated; exp = the base of natural logarithm.

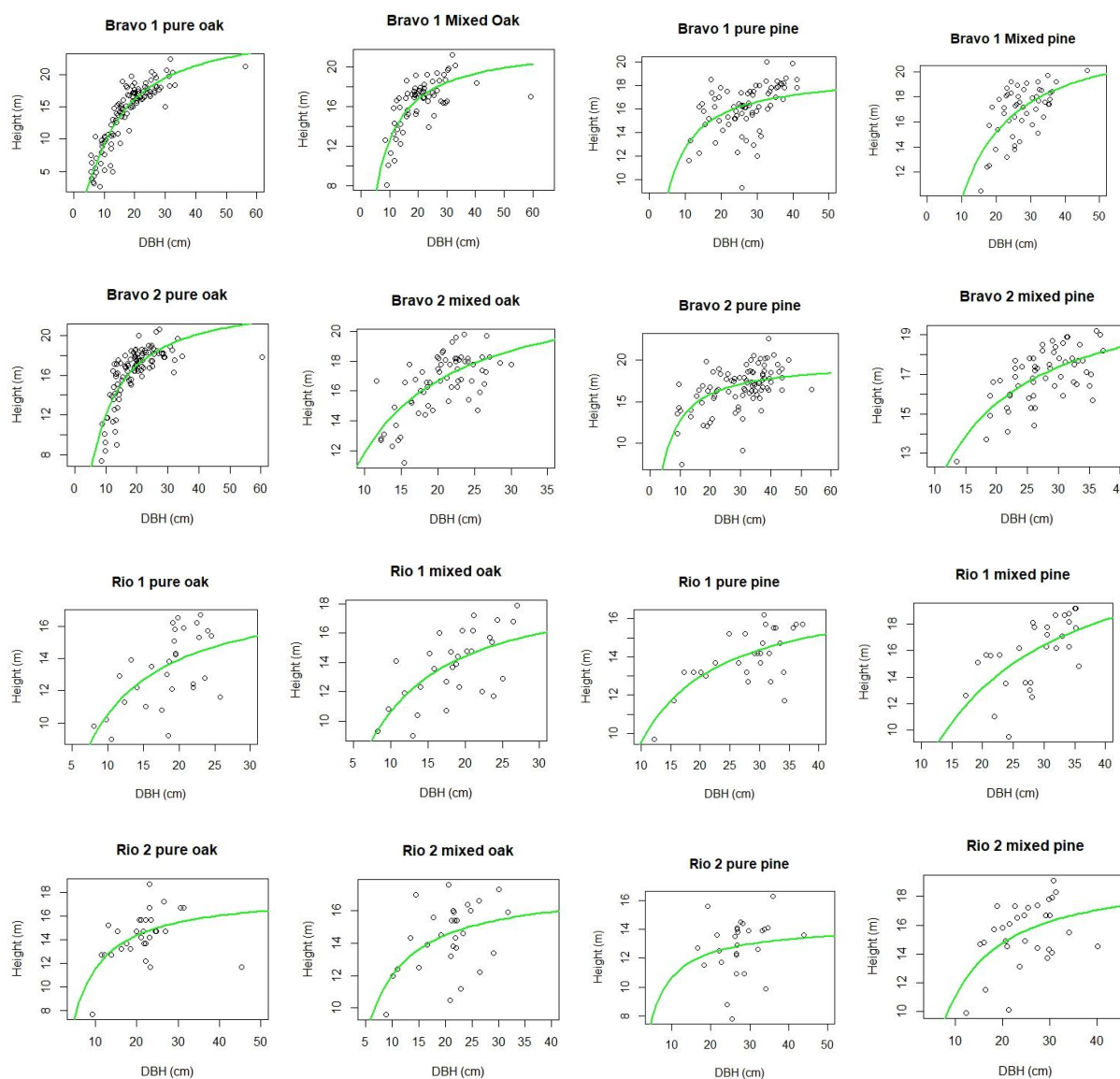


Figure 3: Data and fitted Korf curve model (model no.3) for the different plots in the triplets.

3.3.2. Biomass equation

Beforehand, the dependent and covariate variables were firstly transformed to its natural logarithm in order to observe some statistics assumptions and be fitted as linear models.

- $\ln(\text{AGB})$, natural logarithm of aboveground biomass (AGB);
- $\ln(\text{ST})$, natural logarithm of stem biomass (ST);
- $\ln(\text{TKB})$, natural logarithm of thick branch (TKB);
- $\ln(\text{TNB})$, natural logarithm of thin branch (TNB);
- $\ln(\text{FL})$, natural logarithm of foliage biomass (FL);
- $\ln(\text{Ht})$, natural logarithm of height (Ht);
- $\ln(\text{DBH})$, natural logarithm of diameter at breast height (DBH);

However, $\ln(\text{DBH})$ and $\ln(\text{Ht})$ in the same model developed as independent variables showed low collinearity because the variance inflation factor VIF was less than 5 ($\text{VIF} = 3.15$).

3.3.2.1. Selection of predictors

The dependent or response variable in the regression models was total aboveground biomass and the independent (co-variant) variables were DBH (diameter at breast height or 1.3 m above the ground, cm), Ht (total height, m), combinations of both, crown diameter, crown ratio, and crown base height. However, all the variables were transformed to its logarithmic form before analyzing. The model selection criteria for the independent variables was based on Akaike information criterion (AIC) and Adjusted coefficient of determination (R^2) as statistics and behavior of the models from graphical examination.

3.3.2.2. Biomass allometry of monospecific and mixed stands

In this study, several relations were analyzed to show if there are difference within the monospecific and mixed stands in terms of height-DBH, height-crown base height and height-crown length. The testing of multicollinearity was performed to check that the different variables do not have similar predictive relationship without compromising the accuracy of the model. Moreover, analysis of covariance (ANCOVA) was used to determine mixture effect on tree allometry. Monospecific and mixed stands differences were separated by Tukey's honestly significant difference (HSD) post hoc tests at a significance level of $P < 0.05$. Therefore, the dependent variables are $\ln(\text{AGB})$, $\ln(\text{ST})$, $\ln(\text{TKB})$, $\ln(\text{TNB})$ and finally $\ln(\text{FL})$. All the assumptions of ANCOVA (including normality, homogeneity of variance, homogeneity of regression slopes,

independence, and linearity) were checked.

3.3.2.3. Estimation of total aboveground biomass

Generally, DBH is used in the prediction of aboveground biomass, but combining height and DBH really improves the accuracy of the models (Ruiz-Peinado et al., 2011); Poudel & Temesgen, 2016b). However, graphical exploratory analyses on the relationship between DBH, crown base height, total height, crown length and the dependent variables are needed to better understand the trend of the data. Eight allometric equations (including simple linear, non-linear and mixed effects models) were tested (Table 4). The data for this study was from 2 triplets within 6 plots, thus analysis of the dataset is accurately done by separating variance due to the plots and triplets using a mixed effect model, From Table 1, the non-linear mixed model for the equation no.5 was selected as the best model for predicting total aboveground biomass using DBH and total height for Scots pine and sessile oak. The selection of the final model was based on the parameter significance, regression coefficients and Akaike information criterion (AIC).

Table 12: Biomass models tested for total aboveground biomass for both species

| No | Type | Model |
|----|--------------------|---|
| 1 | Linear | $AGB=b_1*DBH*Ht$ |
| 2 | Linear | $AGB =b_1*DBH$ |
| 3 | Logarithmic | $\ln(AGB) =b_1*\ln(DBH*Ht)$ |
| 4 | Mixed model | $AGB=b_1*(DBH*Ht)$ |
| 5 | Mixed model | $\ln(AGB) =b_1*\ln(DBH)+ b_2*\ln(Ht)$ |
| 6 | Mixed model | $\ln(AGB)=b_1*\ln(DBH*Ht)$ |
| 7 | Mixed model | $\ln(AGB) =b_1*\ln(DBH*Ht)+b_2*\ln(DBH*Ht)^2*Ht$ |
| 8 | Mixed model | $\ln(AGB) =b_1*\ln(DBH*Ht)+b_2*\ln(DBH*Ht)^2+b_3(DBH*Ht)^3$ |

Note: AGB - biomass component (kg); DBH-diameter at breast height (cm); Ht-total height (m); b1, b2 and b3 - model parameters.

3.3.2.4. Estimation of component biomass

For the different tree components, similar equations used in predicting total tree biomass was used in modelling of the biomass. However, Dirichlet regression was used as it observes the additivity. The actual amount of component biomass was obtained by the product of predicted total biomass obtained from model no.5 and predicted proportion of component biomass. Fourteen linear models were used in this study for predicting the biomass in different tree components using Dirichlet regression. Furthermore, Dirichlet regression is much essential with the condition that the component proportion makes a total of 1. It is usually used in modelling of component biomass in Forestry.

3.3.2.5. Evaluation

The performances of the biomass methods were evaluated depending on bias, bias percent, root mean square error and RMSE percent.

$$\begin{aligned} Bias &= \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \\ Bias \% &= 100 \times \frac{Bias}{\bar{Y}} \% \\ RMSE &= \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \\ RMSE \% &= 100 \times \frac{RMSE}{\bar{Y}} \% \end{aligned}$$

where n is the number of trees, y_i and \hat{y}_i are observed and predicted values of AGB or its component, and \bar{Y} is the mean AGB or component biomass.

4. RESULTS

4.1. Stand characteristics

4.1.1. Height and diameter a breast height equation

The full data from both triplet experimental (Palencia -Bravo- and Burgos -Rio) the North of Palencia and the North of Burgos sites were used to build the H-DBH relationship and determine if there are some differences in the monospecific and mixed stands. For the pooled data set of non-linear models (Table 5), Korf Curve was the most appropriate model form selected among the four alternatives because it has the lowest rate of error in terms of residual for predicting height to diameter. Based on these results, The Korf's curve height-diameter models performed better than other models and all parameters of the best fit models were significant ($p \leq 0.005$). However, we can observe from (Table 5) that the parameter estimates values (a and b) are very similar in all the plots for sessile oak (monospecific and mixed), thus sessile oak parameters in monospecific and mixed stands are the same. On the other hand, mixed stands were superior to monospecific stands for Scots pine forests. From anova analysis, the height-diameter relationship was slightly different from mixed and monospecific stands for pine species. (Table 5) shows the fit statistics for both monospecific and mixed stands of Scots pine and sessile oak.

Table 13: Parameter estimates of estimate for equation [3] in Table 3

| Triplet | Species | Trees (n) | a | b |
|----------------|------------|-----------|------------------|------------------|
| Bravo_1 | Pure pine | 74 | 18.9899 (0.7898) | 4.0496 (1.0059) |
| | Pure oak | 114 | 28.2138 (0.9475) | 11.1252 (0.5990) |
| | Mixed_Pine | 51 | 23.4640 (1.4050) | 8.7050 (1.5870) |
| | Mixed_Oak | 65 | 21.0698 (0.7904) | 6.2428 (0.7048) |
| Bravo_2 | Pure pine | 91 | 19.8520 (0.6222) | 4.4045 (0.7814) |
| | Pure oak | 100 | 23.9004 (0.7168) | 6.8575 (0.5430) |
| | Mixed_Pine | 58 | 21.7351 (0.8509) | 6.7246 (1.0389) |
| | Mixed_Oak | 68 | 23.4870 (1.1290) | 6.8330 (0.9540) |
| Rio_1 | Pure pine | 29 | 17.6436 (0.9087) | 6.1607 (1.3455) |
| | Pure oak | 30 | 18.4220 (1.6320) | 5.6020 (1.5300) |
| | Mixed_Pine | 29 | 20.7270 (1.1170) | 7.2470 (1.1660) |
| | Mixed_Oak | 30 | 19.5780 (1.6680) | 6.1070 (1.4870) |
| Rio_2 | Pure pine | 27 | 14.3360 (1.9810) | 2.9650 (3.5990) |
| | Pure oak | 30 | 17.8870 (1.3030) | 4.3800 (1.4140) |
| | Mixed_Pine | 30 | 19.7270 (1.7580) | 5.8380 (2.8130) |
| | Mixed_Oak | 29 | 17.5030 (1.2890) | 3.7700 (1.3800) |

*All estimated parameters were significantly different from 0 (meaning $P < 0.005$).

4.2. Biomass analysis

4.2.1. Variables selection for Aboveground Biomass Estimations

Forward selection of the stepwise regression was performed on the different stands to determine the parameters that best describe total aboveground biomass. However, the total aboveground biomass was highly correlated with diameter at breast height (DBH) and total height (Ht) for pure pine ($R^2 > 0.98$), mixed pine ($R^2 > 0.95$) and pure oak ($R^2 > 0.99$) except for mixed oak whose total aboveground biomass was greatly dependent only on diameter at breast height (DBH) with ($R^2 > 0.98$).

4.2.2. Relationship between independent variables for monospecific and mixed stands

It is important to check out and analyse the relationships between the independent variables (total height, diameter at breast height, crown base height and crown length) to help in the interpretation of the biomass allometric differences in monospecific and mixed stands. However, the relationship

between height and DBH were similar for monospecific and mixed sessile oak stands. There was a difference between the monospecific and mixed Scots pine, but the difference was not significant. Also, the relationship between Height and Crown base height for both stands were unsimilar likewise height and Crown length relationship.

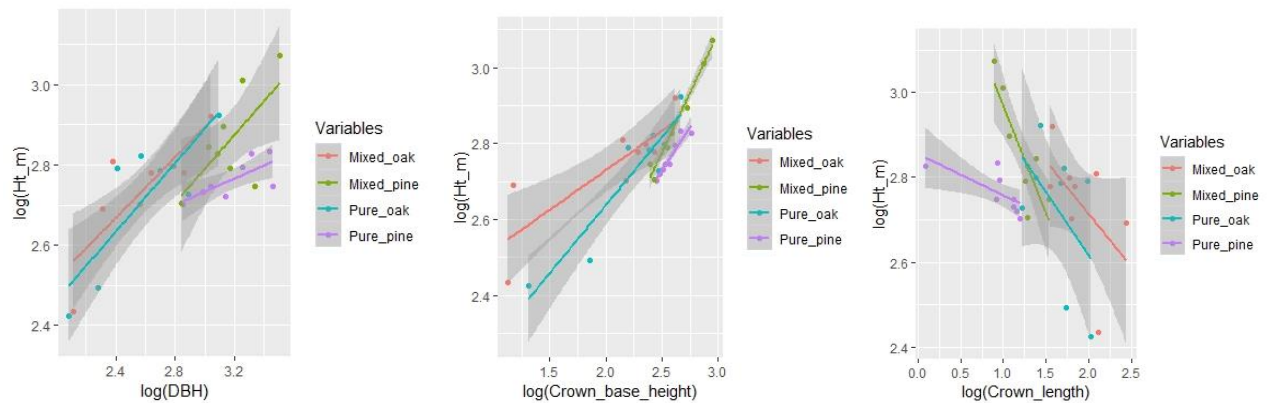


Figure 4: The relationships among independent variables for monospecific and mixed stand trees.

4.2.3. Differences in biomass allometry of monospecific and mixed stands

The determination of the differences in monospecific and mixed stands for pine and oak species was evaluated using $\ln(DBH)$ as covariate. Analysis of covariance, ANCOVA showed no statistical significant differences for the total aboveground biomass, thick branches, and stem in monospecific and mixed Scots pine stand. Thin branches were moderately significant ($p = 0.00186 > 0.01$) and the foliage was slightly significant ($p = 0.01252 > 0.05$). Moreover, there was no difference in total biomass, stem, thin branches, and foliage for both monospecific and mixed oak. Nevertheless, $\ln(TKB)$, $\ln(ST)$ predicted by $\ln(DBH)$ were slightly significant ($p = 0.0399 > 0.05$). Although, ANCOVA results indicate that intercept values of monospecific and pure stands were significantly different for all biomass components.

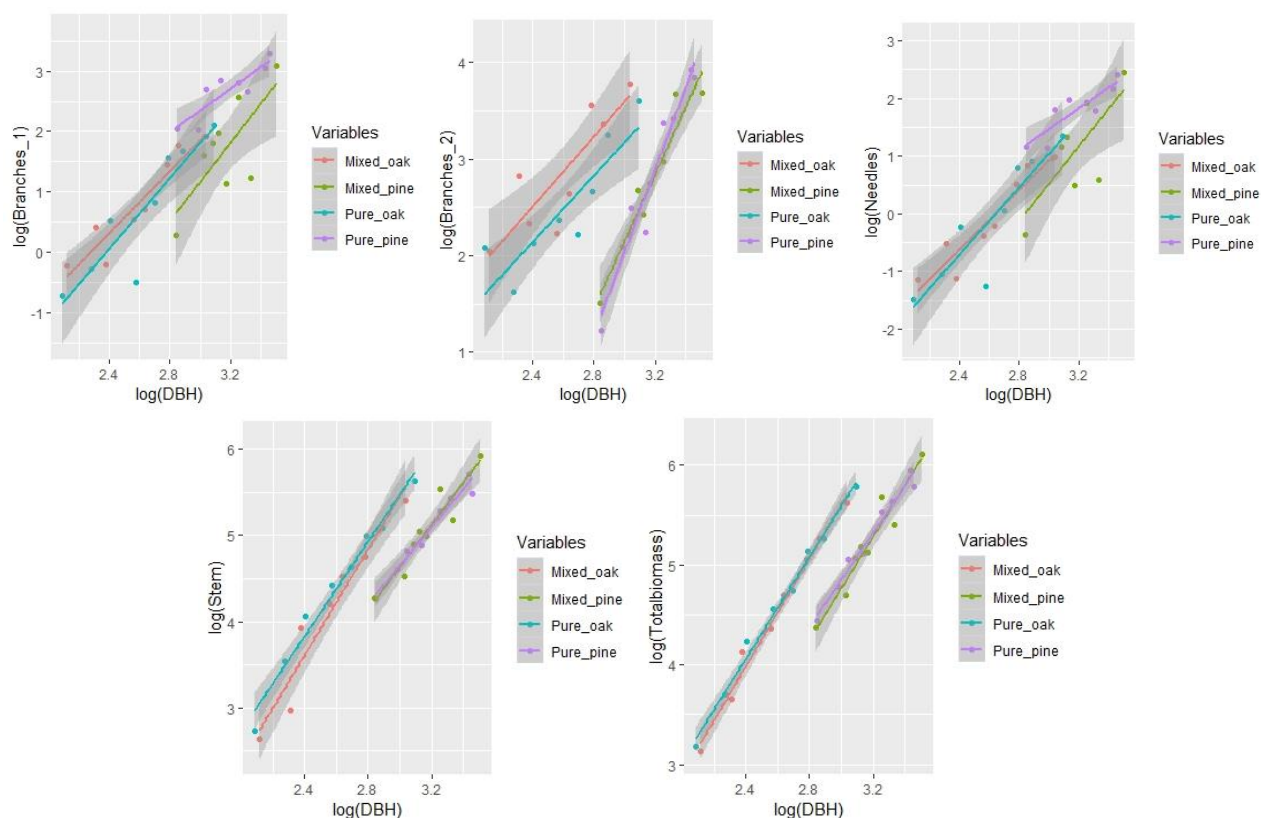


Figure 5: Differences in allometry of monospecific and mixed stand trees

4.2.4. Estimation of biomass component proportions using Dirichlet method

The greater part of the aboveground biomass was found to be in the stem component for Scots pine and sessile oak. The mean biomass proportion of stem with respect to the total above ground biomass (AGB) was 81.8%, whereas thin branch 5.5%, thick branch was 9.6% and foliage was 3.1% in Scots pine. Mean biomass proportion of stem in sessile oak was 78.0%, whereas thin branch and thick branch was 6.8% and 11.1% respectively. The foliage marked 4.1% of the total aboveground in sessile oak. Therefore, foliage was the component that accounted for the smallest proportion in both species. Moreover, to predict the total biomass and proportion biomass, regression procedure was applied, although total biomass and its proportion of various biomass components depend on the species type. For Scots pine, the stem biomass proportion decreases slightly when DBH increases whereas the thick branches, thin branches and foliage increase slightly. On the other hand, the proportion of stem biomass increases for sessile oak and the other component like thick branches, thin branches and foliage decreases slightly (Figure 5).

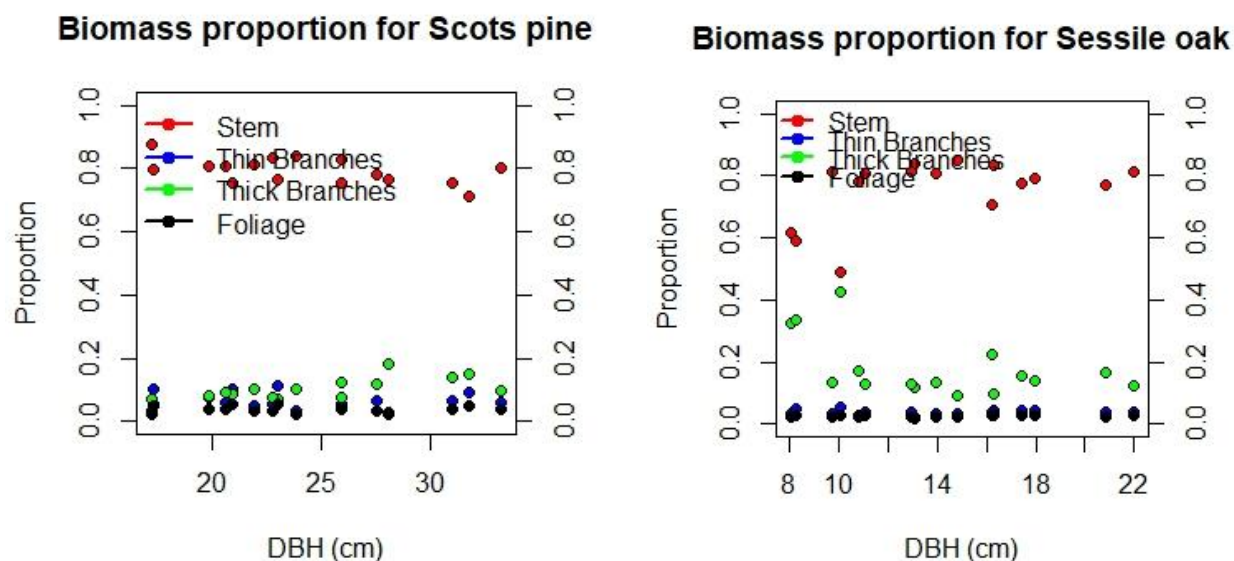


Figure 6: The trend in component biomass proportion with respect with to DBH

4.2.5. Estimation of total aboveground biomass

There is no need to estimate the proportion of different components of biomass from the estimation of the total above-ground biomass as using Dirichlet regression we obtain them directly. Usually the predicted total aboveground biomass facilitates future applications. Non-linear model mixed model fitted by maximum likelihood relating total aboveground biomass with DBH and height (Equation 5) was the best model for both Scots pine and sessile oak. Parameter estimates, standard errors, and log likelihood of this model for both species are shown in (Table 6)

Table 14: Parameter estimates and their standard errors for the In-likelihood model to estimate total aboveground biomass in Scots pine and sessile oak trees.

| Species | Parameter Estimates (S.E) | | | In-likelihood |
|-------------|---------------------------|---------------|---------------|---------------|
| | a | b | c | |
| Scots pine | 0.019 (-0.014) | 2.288 (0.171) | 0.686 (0.307) | -69.973 |
| Sessile oak | 0.052 (0.029) | 2.363 (0.118) | 0.500 (0.281) | -53.933 |

$$(3) \ln(AGB) = a + b * \ln(DBH) + c * \ln(Ht) + e$$

The independent variables were transformed into Logarithmic form because they provided better fit as compared to models in the original form. For Dirichlet regression, the different component proportions are fitted simultaneously, because it ensures that the predicted proportions sum up to 1. The amount of variation explained in Dirichlet regression was high for both species. However, the

statistical evaluation produced by Dirichlet regression are presented in Table 6. The component models were generally as good as the model to fit the total aboveground biomass. The model predicted a higher R^2 for the different component of biomass. Although, the least R^2 for Scots pine was 0.72 in the thin branch and the lowest R^2 for sessile oak was 0.69 for the thin branches as well. (Table 7) indicates clearly that stem biomass proportion for Scots pine was the highest with $R^2=0.88$ but stem biomass proportion had an $R^2=0.81$ which was the highest in sessile oak.

Table 15: Parameter estimates and the standard errors for the component biomass models using the Dirichlet regression where Scots pine includes DBH (cm) tree (Ht, m) and sessile oak includes only Ht (m).

| Model | Parameter Estimates | | | |
|--------------------|---------------------|------|------|-------|
| | Intercept | DBH | Ht | R^2 |
| Scots pine | | | | |
| Stem | -30.13 | 0.10 | 2.05 | 0.76 |
| Thin branch | -32.74 | 0.13 | 2.02 | 0.72 |
| Thick branch | -31.74 | 0.16 | 1.94 | 0.88 |
| Foliage | -33.48 | 0.12 | 2.04 | 0.86 |
| Sessile oak | | | | |
| Stem | -4.70 | - | 0.58 | 0.81 |
| Thin branch | -7.00 | - | 0.57 | 0.69 |
| Thick branch | -6.47 | - | 0.56 | 0.76 |
| Foliage | -7.58 | - | 0.57 | 0.74 |

For Scots pine, RMSEs for stem and thin branch biomass estimation was 3.63% and 15.59% respectively. Also, thick branch was 3.41% and that of foliage was 2.89% while the RMSEs for stem accounted for 7.53%, thin branch was 72.45, thick branch was 5.08% and lastly foliage was 2.3% for sessile oak. Eventually, the stem biomass proportion of R^2 was the highest, it also had the lowest RMSE percent for sessile oak.

Table 16: Evaluation statistics (root mean square error (RMSE), RMSE percent, bias and bias percent) in component biomass estimation in Scot pine and sessile oak trees using Dirichlet regression.

| Component | RMSE (kg) | RMSE percentage | Bias (kg) | Bias percent |
|--------------------|-----------|-----------------|-----------|--------------|
| Scots pine | | | | |
| Stem | 6.38 | 3.63 | 0.21 | 0.12 |
| Thin branch | 4.17 | 15.59 | -3.16 | -11.82 |
| Thick branch | 3.41 | 15.84 | -1.49 | -6.91 |
| Foliage | 2.89 | 54.67 | -5.43 | -102.82 |
| Total | 16.84 | 7.86 | -9.87 | -4.61 |
| Sessile oak | | | | |
| Stem | 7.62 | 7.53 | 0.05 | 0.05 |
| Thin branch | 2.17 | 72.45 | -7.33 | -245.17 |
| Thick branch | 5.08 | 28.60 | -0.69 | -3.88 |
| Foliage | 2.30 | 178.01 | -16.34 | -1266.82 |
| Total | 17.16 | 13.93 | -24.31 | -19.73 |

5. DISCUSSION

The relationship between H and DBH affects strongly the stem biomass allometry. This is because greater height leads to greater stem biomass for the same DBH. Results demonstrated similar patterns of H-DBH relationship for sessile oak in both monospecific and mixed stands. This findings are consistent with other research work on tree allometry (Pretzsch, 2014); Drössler et al., 2015; Dutca et al., 2017). However, the patterns of H-DBH relationship for Scots pine is different in mixed and monospecific stands. We could also observe that the crown base height for both species was higher in mixed than monospecific stands though the crown length of Scots pine and sessile oak in mixture were smaller when compared to monospecific stands (Dieler & Pretzsch, 2013). This may be because the main interaction is greater lateral competition imposed by the other species (e.g. light) as suggested by (Pretzsch et al., 2010; Pretzsch, 2014). Also, this explains how mixed stands will be less vulnerable to fire since mixed stands showed a higher crown base height compared to monospecific stands. It will be difficult for the fire to reach the canopies of the tree, therefore, making mixed stands much more protective against wildfire. The general overview of the results for the component biomass demonstrated no significant differences in the biomass allometry for monospecific and mixed stands for Scots pine except for thin branches (ANCOVA, $p=0.00186$) which was moderately significant and foliage biomass (ANCOVA, $p=0.01252$) which was slightly significant. Sessile oak recorded no difference for total biomass, stem, thin branches and foliage in monospecific and mixed stands. However, the thick branches (ANCOVA, $p=0.0399$) were slightly

significant. The thin branches and needles were significant in mixed Scots pine due to its photosynthesis role played. Therefore, higher total yield in mixed stands with respect to monospecific stands. But thick branches been significant in mixed stands for sessile oak may basically mean lower horizontal competition. The proportion of stem biomass increased slightly with increasing DBH for sessile oak. Similar findings for brutian pine were reported by (De-miguel et al., 2014) Eker et al., 2017). However, in the case for Scots pine is inversely related. Also, the rate of decrease for the stem biomass proportion was lower increasing DBH, in centimeter. It is very important to understand that when the proportion of one component increases, the proportion of the other components decrease as seen in (Figure 3). This could be that Scots pine allocates all its energy on the height rather than DBH because of the competition of light. This confirms with stand dynamics of light-demanding species. However, in the case of sessile oak, this might be that the vertical competition influences the crown biomass as age increases.

Most of the aboveground biomass was found allocated in the stem for both species. The mean biomass proportion of stem with respect to the total above ground biomass (AGB) was 81.8%, whereas thin branch 5.5%, thick branch was also 9.6% and foliage was 3.1% in Scots pine. Mean biomass proportion of stem in sessile oak was 78.0%, whereas thin branch and thick branch was 6.8% and 11.1% respectively. The foliage marked 4.1% of the total aboveground in Sessile oak. Therefore, foliage was the component that accounted for the smallest proportion in both species. Moreover, to predict the total biomass and proportion biomass, regression procedure was applied, although total biomass and its proportion of various biomass components depend on the species type. The proportion of stem to total biomass was again similar in monospecific and mixed stands for both species. Knowing the biomass percentages in the different components is very important to determine which portion of the tree can provide what amount of biomass for different purposes. Dirichlet regression (Poudel & Temesgen, 2016a) was used in the estimation of the proportion of component biomass. One desired property in the component biomass estimation is the property of additivity, which can be attained by simultaneous fitting of component proportions in the Dirichlet regression (Poudel & Temesgen, 2016a).

Generally, biomass allometric models are often developed using dataset from either monospecific or mixed stands. However, these models are used disregarding whether trees grow in monospecific or mixed stands. Results from biomass allometric models predicting total aboveground biomass were not significantly different in monospecific and mixed stands for Scots pine and Sessile oak. The parameter estimates and its standard errors for the non-linear mixed effects model which was used in predicting the total aboveground biomass in (Table 5) explains that the relationship between total aboveground biomass. The aboveground biomass was completely correlated with the predictor variables including DBH and total height. However, stem, thin and thick branches, foliage

biomass were found to be similar for both monospecific and mixed stands using DBH and total height. Biomass equations are very useful to predict total biomass and by compartment so end-user can assess different potential uses as energy production, board industry. In this research work, new set of biomass estimation models for Scots pine and Sessile oak mixed stand. This is the first time such an equation has been built using Dirichlet regression. Most recent research work have use Dirichlet regression because of its superiority over other normal multivariate statistical methods (Dahlhausen et al., 2017). The results also confirms that Dirichlet regression is one of best methods, especially when tree biomass component has similar independent variables (Hijazi & Jernigan, 2009). Therefore, Dirichlet regression is applicable for the biomass estimation since the different tree biomass compartment is influenced by DBH and total height. As to our knowledge it is the first-time biomass equations for pure and mixed stands in sessile oak-Scots pine forests are been assessed. Our research chose DBH and total height as predictors for total biomass except for mixed oak where DBH was the only predictor. Other independent variables like crown ratio, crown length and crown base height were tested for total biomass, but they were not significant. There was no difference in the total biomass allometry of monospecific and mixed stands for both species. This is mainly due to the similarity in the stem component which contributes to majority of the total biomass. Stem biomass allometry was found to be similar for monospecific and mixed stands for Scots pine and sessile oak when using DBH as predictors. As previously found by (Larreta et al., 2017) regarding a very strong relationship between biomass components and total height was determined for oak species, and total height was a significant predictor of all tree biomass components and aboveground biomass for Scots pine species. it is therefore possible to use the biomass equation systems developed for a specific species across different sites in the temperate forests. In all species, most of the biomass was allocated in stem biomass, followed by branches (thick and thin branches), and foliage. The biomass allocation differed between tree components and among species; however, it is necessary to investigate on factors affect the allocation patterns. The proposed species-specific biomass equations can be applied to tree-level data in forest inventories and may also improve the quality of biomass estimates and verify changes in carbon stocks in the temperate forests in the study area. This research should be viewed as part of an ongoing process, and further sampling is required to provide data for a wider range of species with potentially different growth forms and biomass allocation patterns in order to improve our ability to estimate aboveground biomass and carbon for these forests. The method of predicting the biomass proportions were unbiased even though we had limited dataset, it should be tested on a larger sample size.

6. CONCLUSION

The models developed for component biomass estimation from monospecific and mixed stands satisfy the additivity property. Thus, helping in the estimation of biomass for Scots pine and sessile oak. Higher accuracy of the model was obtained by incorporating total height to the model mainly for Scots pine. All the findings for Scots pine and sessile oak of the master's research work demonstrate that species composition has no significant effect on tree biomass allometry, at least for this phase of development and mixing proportions. The covariate relationships, and the proportion of biomass component are similar in monospecific and mixed stands for both species. Tree height has a strong influence on the estimate of aboveground biomass made by allometric models. To improve biomass estimation, the development of generalized height-diameter models that estimate tree height using a subset of randomly sampled trees presents an approach to supplement surveys where only diameters have been measured. Using Dirichlet regression to fit the models, lead to accurate additive biomass equations for monospecific and mixed stands for the two species. These biomass equations for mixed stands could help in accounting for carbon estimation and mitigation against climate change. This study presents the first mixed species allometric equations precisely in Spain. The best model explained, more than 65% of the variation in measured aboveground biomass. It produced the lowest AIC, bias and narrow ranges of errors. This confirms that our model is strong and accurately estimates aboveground biomass. Furthermore, the model has potential for application in other regions, species composition and site characters are almost the same as our study area. Species-specific biomass allocation and allometric equations should be used in other to acquire best estimations, however, when that is not available then we can opt for general biomass allocation and allometric equations. It is hereby recommended that the application of mixed biomass models is transposed to tree biomass in monospecific and vice versa. Future interesting work would be to address the topic of leaf area, canopy biomass partitioning and tree distribution pattern under changing environment conditions. Particularly, biomass modelling for mixed stands for Scots pine and sessile oak should also focus on relating forest variables like tree biomass and volume.

7. ACKNOWLEDGEMENT

I first thank Our Almighty and Merciful God, who has given us the strength and patience to do my research work.

My sincere and greatest thanks go to Dr. Felipe Bravo Oviedo, Professor at the University of Valladolid for the confidence he has given me by agreeing to supervise my master's thesis work. I thank him more for his warm welcome every time I asked for his help and assistance, as well as for his many encouragements, guidance, precious advices and suggestions during my working period.

My warm thanks also go to Dr. Ricardo Ruiz-Peinado, researcher at Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria for his sound advice, constructive remarks and Inistic support that he brought to the completion of my dissertation. May he find here the expression of my best feelings.

I particularly acknowledge the Erasmus Mundus MEDfOR Master program funded by the European Commission. Finally, I would like to thank everyone who has been involved in this work from near and far.

8. REFERENCES

- Agnoletti, M. (2006). Man, forestry, and forest landscapes. Trends and perspectives in the evolution of forestry and woodland history research. *Schweiz. Z. Forstwes.*, (157), 384–392.
- Augusto da Silva Scaranello, M., Alves, L. F., Vieira, S. A., Camargo, P. B. De, Joly, C. A., & Martinelli, L. A. (2012). Height-diameter relationships of tropical Atlantic moist forest trees in southeastern. *Science Agriculture*, 26–37. <https://doi.org/http://dx.doi.org/10.1590/S0103-90162012000100005>
- Batista, J. L., Couto, H. T. Z., & Marquesini, M. (2001). Performance of height-diameter relationship models: analysis in three forest types. *Scientia Forestalis*, 60, 149–163.
- Burquez, A., Chidumayo, E., Savanna, M., & Delitti, W. B. C. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*. <https://doi.org/10.1111/gcb.12629>
- Cañellas, I., Martínez García, F., & Montero, G. (2000). Silviculture and dynamic of *Pinus sylvestris* L. stands in Spain. *Investigación Agraria. Sistemas y Recursos Forestales*, 1, 233–253.
- Coll, L., Ameztegui, A., Collet, C., Löf, M., Mason, B., Pach, M., ... Ponette, Q. (2018). Knowledge gaps about mixed forests: What do European forest managers want to know and what answers can science provide? *Forest Ecology and Management*, (407), 106–115. <https://doi.org/https://doi.org/10.1016/j.foreco.2017.10.055>
- Correia, A. C., Faias, S. P., Ruiz-Peinado, R., Chianucci, F., Cutini, A., Fontes, L., ... Tomé, M. (2018). Generalized biomass equations for Stone pine (*Pinus pinea* L.) across the Mediterranean basin. *Forest Ecology and Management*, 429, 425–436. <https://doi.org/10.1016/j.foreco.2018.07.037>
- Czapowskyj, M. M., Robison, D. J., Briggs, R. D., & White, E. H. (1985). Component biomass equations for black Spruce in Maine. USDA Forest Service, North-Eastern Research Station, Broomall, Pennsylvania, Research p, NE-564.
- Dahlhausen, J., Uhl, E., Heym, M., Biber, P., Ventura, Maurizio Panzacchi, P., Tonon, G., ... Pretzsch, H. (2017). Stand density sensitive biomass functions for young oak trees at four different European sites. *Trees*, 31(6), 1811–1826. <https://doi.org/http://dx.doi.org/10.1007/s00468-017-1586-7>
- Dahlhausen, J., Uhl, E., Heym, M., Biber, P., Ventura, M., Panzacchi, P., ... Pretzsch, H. (2017). Stand density sensitive biomass functions for young oak trees at four different European sites. *Trees*, 31(6), 1811–1826. <https://doi.org/10.1007/s00468-017-1586-7>
- De-miguel, S., Pukkala, T., Assaf, N., & Shater, Z. (2014). Intra-specific differences in allometric equations for aboveground biomass of eastern Mediterranean *Pinus brutia*. *Annals of Forest*

- Science, 71(1), 101–112. <https://doi.org/10.1007/s13595-013-0334-4>
- del Río, M., Pretzsch, H., Alberdi, I., Bielak, K., Bravo, F., Brunner, A., ... Bravo-Oviedo, A. (2018). Characterization of the structure, dynamics, and productivity of mixed-species stands. *European Journal of Forest Research*, 135, 27–71.
- del Rio, M., & Sterba, H. (2009). Comparing volume growth in pure and mixed stands of *Pinus sylvestris* and *Quercus pyrenaica*. *Annals of Forest Science*. <https://doi.org/10.1051/forest/2009035>
- Dieler, J., & Pretzsch, H. (2013). Morphological plasticity of European beech (*Fagus sylvatica* L.) in pure and mixed-species stands. *Forest Ecology and Management*, 295, 97–108. <https://doi.org/10.1016/j.foreco.2012.12.049>
- Drössler, L., Övergaard, R., Ekö, P. M., Gemmel, P., & Böhlenius, H. (2015). Early development of pure and mixed tree species plantations in Snogeholm, Southern Sweden. *Scandinavian Journal of Forest Research*, 30, 304–316.
- Dutca, I., Mather, R., & Ioras, F. (2017). Tree biomass allometry during the early growth of Norway spruce (*Picea abies*) varies between pure stands and mixtures with European beech (*Fagus sylvatica*). *Canadian Journal of Forest Research*, 48, 77–84.
- Eker, M., Poudel, K. P., & Özçelik, R. (2017). Aboveground biomass equations for small trees of brutian pine in Turkey to facilitate harvesting and management. *Forests*, 8(12), 12–13. <https://doi.org/10.3390/f8120477>
- Forrester, D. I., & Pretzsch, H. (2015). On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. *Tamm Review: Forest Ecology and Management*, 356, 41–53.
- Gamfeldt, L., Snäll, T., Bagchi, R., Jonsson, M., Gustafsson, L., Kjellander, P., ... Bengtsson, J. (2013). Higher levels of multiple ecosystem services are found in forests with more tree species. *Nature Communications*, 4(May 2013). <https://doi.org/10.1038/ncomms2328>
- Gutiérrez, O. A., & Plaza, P. F. (1967). *Características físico-mecánicas de las maderas españolas*. Ministerio de Agricultura, Madrid, 103.
- Henry, M., Picard, N., Trotta, C., Manlay, R. J., Valentini, R., Bernoux, M., & Saint-André, L. (2011). Estimating tree biomass of sub-Saharan African forests: A review of available allometric equations. *Silva Fennica*, 45(3), 477–569. <https://doi.org/10.14214/sf.38>
- Heym, M., Ruíz-Peinado, R., del Río, M., Bielak, K., Forrester, D. I., Dirnberger, G., ... Pretzsch, H. (2017). EuMIXFOR empirical forest mensuration and ring width data from pure and mixed stands of Scots pine (*Pinus sylvestris* L.) and European beech (*Fagus sylvatica* L.) through Europe. *Annals of Forest Science*, 74(3), 9. <https://doi.org/10.1007/s13595-017-0660-z>
- Hijazi, R. H., & Jernigan, R. W. (2009). Modelling compositional data using Dirichlet regression models. *Journal of Applied Probability and Statistics*, 4, 77–91.

- Huang, S., Titus, S. J., & Wiens, D. P. (1992). Comparison of nonlinear height–diameter functions for major Alberta tree species. *Canadian Journal of Forest Research*, 22(9), 1297–1304. <https://doi.org/https://doi.org/10.1139/x92-172>
- Hunter, M. O., Keller, M., Victoria, D., & Morton, D. C. (2013). Tree height and tropical forest biomass estimation. *Biogeosciences*, 8385–8399. <https://doi.org/10.5194/bg-10-8385-2013>
- Karlik, J. ., & Chojnacky, D. (2014). Biomass and carbon data from blue oaks in a California oak savanna. *Biogeosciences*, (62), 228±232.
- Larreta, B. V., López-Sánchez, C. A., Corral-Rivas, J. J., López-Martínez, J. O., Aguirre-Calderón, C. G., & Álvarez-González, J. G. (2017). Allometric Equations for Estimating Biomass and Carbon Stocks in the Temperate Forests of North-Western Mexico. *Forests*, 8(8), 1–20. <https://doi.org/10.3390/f8080269>
- Légaré, S., Paré, D., & Bergeron, Y. (2004). The responses of black spruce growth to an increased proportion of aspen in mixed stands. *Canadian Journal of Forest Research*, (34), 405–416.
- Lu, H., Mohren, G. M. J., Schelhaas, M., Bouwman, M., & Sterck, F. J. (2018). Species Mixing Effects on Forest Productivity: A Case Study at Stand-, Species- and Tree-Level in the Netherlands. *Forests*, 1, 1–21. <https://doi.org/10.3390/f9110713>
- Luo, Y., Wang, X., Ouyang, Z., Lu, F., Feng, L., & Tao, J. (2019). ChinAllomeTree 1.0: China's normalized tree biomass equation dataset, 1–75.
- Maier, M. J. (n.d.). DirichletReg: Dirichlet Regression for Compositional Data in R. Research Report Series/Department of Statistics and Mathematics, 125. WU Vienna University of Economics and Business, Vienna, 1, 1–13. Retrieved from <http://epub.wu.ac.at/4077/>
- Mason, W. L., & Alía, R. (2000). Current and future status of Scots pine (*Pinus Sylvestris* L.) forests in Europe. *Investigación Agraria. Sistemas y Recursos Forestales*, 9, 317–335.
- Montero, G., Ruiz-Peinado, R., & Muñoz, M. (2005). Producción de biomasa y fijación de CO₂ por los bosques españoles. *Monografías INIA Serie Forestal*, 270.
- Mosseler, A., Major, J., Labrecque, M., & Larocque, G. (2014). Allometric relationships in coppice biomass production for two North American willows (*Salix* spp.) across three different sites. *Forest Ecology and Management*, (320), 190±196.
- Næsset, E. (2004). Estimation of above-and below-ground biomass in boreal forest ecosystems. *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, 36(8/W2), 145–148. <https://doi.org/10.1007/bf00835333>
- Parresol, B. R. (1999). Assessing Tree and Stand Biomass: Critical Comparisons. *Forest Science*, 45(4).
- Parresol, B. R. (2001). Additivity of nonlinear biomass equations. *Forest Science*, (Parresol).
- Pastor, J., Aber, J., & Melillo, J. (1984). Biomass prediction using generalized allometric regressions for some northeast tree species. *Forest Ecology and Management*, 7, 265±274.

- Poudel, K. P., & Temesgen, H. (2015). Methods for estimating aboveground biomass and its components for Douglas-fir and lodgepole pine trees. *Canadian Journal of Forest Research*, 46(1), 77–87. <https://doi.org/10.1139/cjfr-2015-0256>
- Poudel, K. P., & Temesgen, H. (2016a). Developing biomass equations for Western hemlock and red alder trees in Western Oregon forests. *Forests*, 7(4). <https://doi.org/10.3390/f7040088>
- Poudel, K. P., & Temesgen, H. (2016b). Methods for estimating aboveground biomass and its components for Douglas-fir and lodgepole pine trees. *Canadian Journal of Forest Research*, 46(1), 77–87. <https://doi.org/10.1139/cjfr-2015-0256>
- Pretzsch, H. (2014). Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. *Forest Ecology and Management*, (327), 251–264. <https://doi.org/https://doi.org/10.1016/j.foreco.2014.04.027>
- Pretzsch, H. (2018). Growth and Structure in Mixed-species stands compared with monocultures: Review and perspectives. In: Bravo-Oviedo A, del Rio M, Pretzsch H (eds) *Dynamics, silviculture and management of mixed forests*. Springer-Verlag, 131–183.
- Pretzsch, H., Block, J., Dieler, J., Dong, P. H., Nagel, J., Spellmann, H., ... Andreas, Z. (2010). Comparison between the productivity of pure and mixed stands of Norway spruce and European beech along an ecological gradient. *Annals of Forest Science*, 67(7), 712–712. <https://doi.org/https://doi.org/10.1051/forest/2010037>
- Pretzsch, H., Forrester, D. I., & Rötzer, T. (2015). Representation of species mixing in forest growth models: A review and perspective. *Ecological Modelling*, 313, 276–292. <https://doi.org/10.1016/j.ecolmodel.2015.06.044>
- Ruiz-Peinado, R., del Rio, M., & Montero, G. (2011). New models for estimating the carbon sink capacity of Spanish softwood species. *Forest Systems*, 20(1), 176. <https://doi.org/10.5424/fs/2011201-11643>
- Ruiz-Peinado, R., Montero, G., & del Rio, M. (2012). Biomass models to estimate carbon stocks for hardwood tree species. *Forest Systems*, 21(1), 42–52. <https://doi.org/10.5424/fs/2112211-02193>
- Stanners, D., & Bourdeau, P. (1995). Landscapes – Chapter 8. In *Europe's Environment – The Dobris Assessment*. European Environment Agency.
- Vashum, K. T., & Jayakumar, S. (2012). Ecosystem & Ecography Methods to Estimate Above-Ground Biomass and Carbon Stock in Natural Forests - A Review. *J. Ecosyst. Ecogr.* 2(4), (September), 1–7. <https://doi.org/10.4172/2157-7625.1000116>
- Wang, X., Fang, J., & Zhu B. (2008). Forest biomass and root±shoot allocation in northeast China. *Forest Ecology and Management*, (255), 4007±4020.
- Xiao, C., Yuste, J. C., Janssens, I. A., Roskams, P., Nachtergale, L., Carrara, A., ... Ceulemans, R. (2003). Above- and belowground biomass and net primary production in a 73-year-old Scots

pine forest Above- and belowground biomass and net primary production in a 73-year-old Scots pine forest. *Tree Physiology*, 23(8), 505–516. <https://doi.org/10.1093/treephys/23.8.505>

Yuen, J. Q., Ziegler, A. D., Webb, E. L., & Ryan, C. M. (2013). Forest Ecology and Management Uncertainty in below-ground carbon biomass for major land covers in Southeast Asia. *Forest Ecology and Management*, 310, 915–926. <https://doi.org/10.1016/j.foreco.2013.09.042>

Zellner, A., & Theil, H. (1962). Three-stage least squares: simultaneous estimation of simultaneous equations. *Econometrica*, (30), 54–78.

Zellner, Arnold. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, 57(298), 348–368.

Zianis, D., Muukkonen, P., Mäkipää, R., & Mencuccini, M. (2005). Biomass and stem volume equations for tree species in Europe. *Silva Fennica Monographs*, 4, 63.

Ziegler, A. D., Phelps, J., Yuen, J. I. A. Q. I., & Webb, E. L. (2012). Carbon outcomes of major land-cover transitions in SE Asia: great uncertainties and REDD + policy implications, 1–13. <https://doi.org/10.1111/j.1365-2486.2012.02747.x>

ANNEX

1. R scripts

3.1. Dirichlet regression

```
#####
##### #
# Dirichelet regression modelling program for estimatebiomass proportion for both species #
#####
#####

## SCRIPT for DIRICHLET REGRESSION

#   Dirichelet regression modelling program to estimate biomass proportion

## RESULTS FROM THIS FITTING ARE PERCENTAGE PARAMETERS FOR THE DIFFERENT
BIOMASS COMPONENTS

## THEY SHOULD BE APPLIED TO TOTAL BIOMASS ESTIMATIONS

# Clean behind run history

rm(list=ls())
```

```

# Load working directory

setwd("C:/MEDfOR_Spain/MScThesis_MEDFOR_UVa/Data/First_Work_INIA/Eric
Cudjoe_Palencia")

# Read data

datos<-read.csv2('Thesis_Work_2.csv', sep=',',dec='.',header=T, na.string='NA')

## Selection of required data

# Selection of required data

## PINE DATA IN THIS CASE --> change to oak data for the oak biomass models

datos_dirichlet<-subset(datos,datos$Variables=="Mixed pine"|datos$Variables=="Pure pine")

# Pine biomass data: total biomass, Stem biomass, branch biomass, needle biomass

# Install packages

packs<-c('Formula', 'rgl', 'DirichletReg', 'colorspace')

sapply(packs,require,character.only=TRUE)

# Introduce mean of biomass (dependent) variable by biomasstotal (to calculate precision
parameter)

datos_dirichlet$Stemp=with(datos_dirichlet,(Stem/Totalbiomass))

datos_dirichlet$Branches_1_1p=with(datos_dirichlet,(Branches_1/Totalbiomass))

datos_dirichlet$Branches_2_1p=with(datos_dirichlet,(Branches_2/Totalbiomass))

datos_dirichlet$Needlesp=with(datos_dirichlet,(Needles/Totalbiomass))

head(datos_dirichlet)

# Make format data for Dirichlet's Regression

# This function prepares a matrix with compositional variables for further processing #12:15 values
to transform

AL <- DR_data(datos_dirichlet[, 12:15],trafo=TRUE)

AL

print(AL,type=c("processed"))

attr(AL,"Y.original")

attr(AL,"transformed")

attr(AL,"normalized")

```

```

attr(AL,"valid_obs")

attr(AL,"obs")

attr(AL,"dim.names")

# Some figures for biomass percentages

plot(AL, main ='Biomass percentage', cex=.16,a2d=list(colored=TRUE, c.grid=TRUE,trafo=TRUE))

##### OR #####

plot(AL, cex = 0.16, a2d = list(colored = FALSE, c.grid = FALSE))

plot(rep(datos_dirichlet$DBH,4), as.numeric(AL), ylim=0:1, pch=21,

      bg=rep(c("red","blue","green","black"), each=16),

      xlab="DBH (cm)", ylab="Biomass proportion")

legend("topleft", legend = c("Stem", "Thin Branch", "Thick Branch", "Foliage"), lwd = 2, col =
c("red","blue","green","black"),

      pt.bg = c("red","blue","green","black"), pch = 21, bty = "n")

# Make format data for Dirichlet's Regression

# Null model (only intercept)

model1<- DirichReg(AL ~ 1, datos_dirichlet)

model1

coef(model1)

summary(model1)

anova(model1)

# Different combination with DBH and Ht_m for fitting commom models

model2 <- DirichReg(AL ~ DBH,datos_dirichlet)

coef(model2)

summary(model2)

predict(model2, newdata = data.frame("DBH" = seq(10, 100, 1000)))

residuals(model2,type=c('standardized'))

anova(model2)

model3 <- DirichReg(AL ~ DBH+(Ht_m), datos_dirichlet)

```

```

coef(model3)
summary(model3)
anova(model3)
residuals(model3,type=c('standardized'))
anova(model3
model4<- DirichReg(AL ~ (DBH)+I((DBH)^2), datos_dirichlet)
summary(model4)
anova(model4)
model5<- DirichReg(AL ~ I(DBH^2)+Ht_m, datos_dirichlet)
summary(model5)
anova(model5)
model6<- DirichReg(AL ~ DBH^2+Ht_m^2, datos_dirichlet)
coef(model6)
summary(model6)
predict(model6,newdata=data.frame("DBH"=seq(10,10,1000), "Ht_m"=seq(10,10,100))
residuals(model6,type=c('standardized'))
anova(model6)
model7<- DirichReg(AL ~ I((DBH+Ht_m)^2), datos_dirichlet)
summary(model7)
anova(model7)
model8<- DirichReg(AL ~ I((DBH)^2), datos_dirichlet)
summary(model8)
anova(model8)
model9 <- DirichReg(AL ~ DBH+I((DBH)^2), datos_dirichlet)
coef(model9)
summary(model9)
anova(model9)
model10 <- DirichReg(AL ~ (Ht_m)+I((Ht_m)^2), datos_dirichlet)

```



```

coef(model10)
summary(model10)
anova(model10)
model11 <- DirichReg(AL ~ I(DBH*Ht_m), datos_dirichlet)
summary(model11)
anova(model11)
model12 <- DirichReg(AL ~ I((DBH)^2*Ht_m), datos_dirichlet)
summary(model12)
anova(model12)
# Fit quadratic Dirichlet regression models ("common")
model13 <- DirichReg(AL ~ I(DBH+Ht_m) + I((DBH+Ht_m)^2), data = datos_dirichlet)
summary(model13)
anova(model13)
model14 <- DirichReg(AL ~ I((DBH^2*Ht_m)) + I((DBH^2*Ht_m)), data = datos_dirichlet)
summary(model14)
anova(model14)
model15 <- DirichReg(AL ~ I(DBH+Ht_m) + I((DBH^2*Ht_m^2)), data = datos_dirichlet)
coef(model15)
summary(model15)
anova(model15)
# Looking for the better model via parameter significance...
#### You can also look for those better models and also those that showed the lower AIC value
## Caution!!!! In order to select the best model, if there are some models that showed problems
(no convergence)
## the names of the models (model1, model2,...) that appears in the next section are not the right
number for
## the models that we have run previously. Models are presented with a continuous number.

```

```

anova(model1, model2, model3, model4, model5, model6, model7, model8, model9, model10,
model11, model12, model13, model14, model15)

## Now, we can check the biological behaviour of the selected models

## First example with two variables '(DBH)+(Ht_m)'

## MODEL 3

par(mar = c(4, 4, 4, 4) + 0.1)

plot(rep(datos_dirichlet$DBH, 4), as.numeric(AL), pch = 21, bg = rep(c("red","blue","green","black"),
each = 16), xlab = "DBH (cm)", ylab = "Proportion", ylim = 0:1, main = "Biomass
proportion. Model 3")

Xnew3 <- data.frame(DBH = seq(min(datos_dirichlet$DBH), max(datos_dirichlet$DBH),length.out =
117),Ht_m=seq(min(datos_dirichlet$Ht_m), max(datos_dirichlet$Ht_m),length.out = 117))

for (i in 1:4) lines(cbind(Xnew3, predict(model3, Xnew3)[, i]), col = c("red","blue","green","black")[i],
lwd = 2)

legend("topleft", legend = c("Stem","Branches_1", "Branches_2", "Needles" ), lwd = 2, col =
c("red","blue","green","black"),

pt.bg = c("red","blue","green","black"), pch = 21, bty = "n")

par(new=TRUE)

plot(cbind(Xnew3, predict(model3, Xnew3, F, F, T)), lty = "24", type = "l", ylim = c(0,
max(predict(model3, Xnew3, F, F, T))), axes = F, ann = F, lwd = 2)

axis(4)

mtext(expression(paste("Precision (", phi, ")", sep = "")), 4, line = 3)

legend("top", legend = c(expression(hat(mu)[c] == hat(alpha)[c]/hat(alpha)[0])),
expression(hat(phi) == hat(alpha)[0])), lty = c(1, 2), lwd = c(3, 2), bty = "n")

#####.....#####

## TRY FOR MODEL 15 example with two variables '(DBH)+(Ht_m)'

## MODEL 15

par(mar = c(4, 4, 4, 4) + 0.1)

plot(rep(datos_dirichlet$DBH, 4), as.numeric(AL), pch = 21, bg = rep(c("red","blue","green","black"),
each = 16), xlab = "DBH (cm)", ylab = "Proportion",

ylim = 0:1, main = "Biomass proportion. Model 16")

```

```

Xnew16 <- data.frame(DBH = seq(min(datos_dirichlet$DBH), max(datos_dirichlet$DBH),length.out
= 40),Ht_m=seq(min(datos_dirichlet$Ht_m), max(datos_dirichlet$Ht_m),length.out = 40))

for (i in 1:4) lines(cbind(Xnew16, predict(model16, Xnew16)[, i]), col =
c("red","blue","green","black")[i], lwd = 2)

legend("topleft", legend = c("Stem","Branches_1", "Branches_2", "Needles" ), lwd = 2, col =
c("red","blue","green","black"),

      pt.bg = c("red","blue","green","black"), pch = 21, bty = "n")

par(new=TRUE)

plot(cbind(Xnew16, predict(model16, Xnew16, F, F, T)), lty = "24", type = "l", ylim = c(0,
      max(predict(model16, Xnew16, F, F, T))), axes = F, ann = F, lwd = 2)

axis(4)

mtext(expression(paste("Precision (", phi, ")", sep = "")), 4, line = 3)

legend("top", legend = c(expression(hat(mu[c] == hat(alpha)[c]/hat(alpha)[0])),
      expression(hat(phi) == hat(alpha)[0])), lty = c(1, 2), lwd = c(3, 2), bty = "n")

#####

## Example with two variables (DBH^2)+Ht_m

#'MODEL 5'

par(mar = c(4, 4, 4, 4) + 0.1)

plot(rep(datos_dirichlet$DBH, 4), as.numeric(AL), pch = 21, bg = rep(c("red","blue","green","black"),
      each = 16), xlab = "DBH (cm)", ylab = "Proportion", ylim = 0:1, main = "Biomass
proportion. Model 4")

Xnew4 <- data.frame(DBH = seq(min(datos_dirichlet$DBH), max(datos_dirichlet$DBH),length.out =
117))

for (i in 1:4) lines(cbind(Xnew4, predict(model4, Xnew4)[, i]), col = c("red","blue","green","black")[i],
lwd = 2)

legend("topleft", legend = c("Stem","Branches_1", "Branches_2", "Needles" ), lwd = 2, col =
cc("red","blue","green","black"),

      pt.bg = c("red","blue","green","black"), pch = 21, bty = "n")

par(new=TRUE)

plot(cbind(Xnew4, predict(model4, Xnew4, F, F, T)), lty = "24", type = "l", ylim = c(0,
      max(predict(model4, Xnew4, F, F, T))), axes = F, ann = F, lwd = 2)

```

```

axis(4)

mtext(expression(paste("Precision (", phi, ")", sep = "")), 4, line = 3)

legend("top", legend = c(expression(hat(mu[c] == hat(alpha)[c]/hat(alpha)[0])),
                           expression(hat(phi) == hat(alpha)[0])), lty = c(1, 2), lwd = c(3, 2), bty = "n")

## Figure. Model6 with variable DBH^2 'MODEL6'

par(mar = c(4, 4, 4, 4) + 0.1)

plot(rep(datos_dirichlet$DBH, 4), as.numeric(AL), pch = 21, bg = rep(c("red", "blue", "green", "black")
                           , each = 16), xlab = "DBH (cm)", ylab = "Proportion", ylim = 0:1, main = "Biomass proportion
for Scots pine")

Xnew6 <- data.frame(DBH = seq(min(datos_dirichlet$DBH), max(datos_dirichlet$DBH), length.out =
117))

for (i in 1:3) lines(cbind(Xnew6, predict(model6, Xnew6)[, i]), col = c("red", "blue", "green", "black")[i],
lwd = 2)

legend("topleft", legend = c("Stem", "Thin Branches", "Thick Branches", "Foliage"), lwd = 2, col =
c("red", "blue", "green", "black"),

      pt.bg = c("red", "blue", "green", "black"), pch = 21, bty = "n")

par(new=TRUE)

plot(cbind(Xnew3, predict(model6, Xnew6, F, F, T)), lty = "24", type = "l", ylim = c(0,
max(predict(model6, Xnew6, F, F, T))), axes = F, ann = F, lwd = 2)

axis(4)

mtext(expression(paste("Precision (", phi, ")", sep = "")), 4, line = 3)

legend("top", legend = c(expression(hat(mu[c] == hat(alpha)[c]/hat(alpha)[0])),
                           expression(hat(phi) == hat(alpha)[0])), lty = c(1, 2), lwd = c(3, 2), bty = "n")

#####

## Figure. Model 10 with variable DBH^2 #

#'MODEL10'

par(mar = c(4, 4, 4, 4) + 0.1)

plot(rep(datos_dirichlet$DBH, 4), as.numeric(AL), pch = 21, bg = rep(c("red", "blue", "green",
"black"),

```

```

each = 16), xlab = "DBH (cm)", ylab = "Proportion", ylim = 0:1, main = "Biomass
proportion. Model 10")

Xnew10 <- data.frame(DBH = seq(min(datos_dirichlet$DBH), max(datos_dirichlet$DBH),length.out
= 117),Ht_m=seq(min(datos_dirichlet$Ht_m), max(datos_dirichlet$Ht_m),length.out = 117))

for (i in 1:3) lines(cbind(Xnew10, predict(model10, Xnew10)[, i]), col =
c("red","blue","green","black")[i], lwd = 2)

legend("topleft", legend = c("Stem","Branches_1", "Branches_2", "Needles" ), lwd = 2, col =
c("red","blue","green","black"),

pt.bg = c("red","blue","green","black"), pch = 21, bty = "n")

par(new=TRUE)

plot(cbind(Xnew10, predict(model10, Xnew10, F, F, T)), lty = "24", type = "l", ylim = c(0,
max(predict(model10, Xnew10, F, F, T))), axes = F, ann = F, lwd = 2)

axis(4)

mtext(expression(paste("Precision (", phi, ")", sep = "")), 4, line = 3)

legend("top", legend = c(expression(hat(mu[c] == hat(alpha)[c]/hat(alpha)[0])),
expression(hat(phi) == hat(alpha)[0])), lty = c(1, 2), lwd = c(3, 2), bty = "n")

#####

## Figure. Model 16 with variable DBH^2 .

#MODEL 16'

par(mar = c(4, 4, 4, 4) + 0.1)

plot(rep(datos_dirichlet$DBH, 4), as.numeric(AL), pch = 21, bg=rep(c("red","blue","green", "pink"),
each = 16), xlab = "dbh (cm)", ylab = "Proportion", ylim = 0:1, main = "Biomass
proportion. Model 16")

Xnew16 <- data.frame(DBH = seq(min(datos_dirichlet$DBH), max(datos_dirichlet$DBH),length.out
= 117))

for (i in 1:3) lines(cbind(Xnew16, predict(model16, Xnew16)[, i]), col = c("red","blue","green",
"pink")[i], lwd = 2)

legend("topleft", legend = c("Branches_1", "Branches_2", "Needles", "Stem"), lwd = 2, col = c("red",
"blue", "green", "pink"), pt.bg = c("red", "blue", "green", "pink"), pch =
21, bty = "n")

par(new=TRUE)

```

```

plot(cbind(Xnew16, predict(model16, Xnew16, F, F, T)), lty = "24", type = "l", ylim = c(0,
                                     max(predict(model16, Xnew16, F, F, T))), axes = F, ann = F, lwd = 2)

axis(4)

mtext(expression(paste("Precision (", phi, ")", sep = "")), 4, line = 3)

legend("top", legend = c(expression(hat(mu)[c] == hat(alpha)[c]/hat(alpha)[0])),
      expression(hat(phi) == hat(alpha)[0])), lty = c(1, 2), lwd = c(3, 2), bty = "n")

## Statistic calculations and graphics

#MODEL 3 -> Model 3: DirichReg(formula = AL ~ DBH + (Ht_m), data = datos_dirichlet)

summary(model3)

fitted(model3)

predict(model3,newdata=data.frame("DBH"=seq(10,10,1000), "Ht_m"=seq(10,10,100)))

residuals(model3,type=c('standardized'))

confint(model3)

confint(model3, exp = TRUE)

logLik(model3)

round(vcov(model3), 3)

#Compute predicted(fitted) percentages of component biomass

datos_dirichlet$Stempp_m3 <- (fitted(model3)[,"Stemp"])

datos_dirichlet$Branches_1pp_m3 <- (fitted(model3)[,"Branches_1_1p"])

datos_dirichlet$Branches_2pp_m3 <- (fitted(model3)[,"Branches_2_1p"])

datos_dirichlet$Needlespp_m3 <- (fitted(model3)[,"Needlesp"])

#Compute predicted(fitted) of component biomass

datos_dirichlet$pStem_m3 <- with(datos_dirichlet,(Stempp_m3 *Totalbiomass))

datos_dirichlet$pBranches_1_m3 <-with(datos_dirichlet,(Branches_1pp_m3 *Totalbiomass))

datos_dirichlet$pbranches_2_m3 <-with(datos_dirichlet,(Branches_2pp_m3 *Totalbiomass))

datos_dirichlet$pNeedles_m3 <-with(datos_dirichlet,(Needlespp_m3 *Totalbiomass))

datos_dirichlet$pbiomastotal_m3 <-with(datos_dirichlet,(pStem_m3 +pBranches_1_m3
+pbranches_2_m3 +pNeedles_m3 ))

```

```

# Compute mean bias of component biomass

# -<caution 16 is the number of observations (n)

datos_dirichlet$Stembias_m3=with(datos_dirichlet,((100/16)*(Stem-pStem_m3)/Stem))

mean(datos_dirichlet$Stembias_m3)

datos_dirichlet$Branches_1bias_m3=with(datos_dirichlet,((100/16)*(Branches_1-
pBranches_1_m3)/Branches_1))

mean(datos_dirichlet$Branches_1bias_m3)

datos_dirichlet$Branches_2bias_m3=with(datos_dirichlet,((100/16)*(Branches_2-
pbranches_2_m3)/Branches_2))

mean(datos_dirichlet$Branches_2bias_m3)

datos_dirichlet$Needlesbias_m3=with(datos_dirichlet,((100/16)*(Needles-pNeedles_m3)/Needles))

mean(datos_dirichlet$Needlesbias_m3)

datos_dirichlet$totalbias_m3<-
with(datos_dirichlet,(Stembias_m3+Branches_1bias_m3+Branches_2bias_m3+Needlesbias_m3))

mean(datos_dirichlet$totalbias_m3)

#Compute mean RMSE of component biomass

datos_dirichlet$Stermrmse_m3<-with(datos_dirichlet,sqrt((Stem-pStem_m3)^2)/16)

sum(datos_dirichlet$Stermrmse_m3)

datos_dirichlet$Branches_1rmse_m3<-with(datos_dirichlet,sqrt((Branches_1-
pBranches_1_m3)^2)/16)

sum(datos_dirichlet$Branches_1rmse_m3)

datos_dirichlet$Branches_2rmse_m3<-with(datos_dirichlet,sqrt((Branches_2-
pbranches_2_m3)^2)/16)

sum(datos_dirichlet$Branches_2rmse_m3)

datos_dirichlet$Needlesrmse_m3<-with(datos_dirichlet,sqrt((Needles-pNeedles_m3)^2)/16)

sum(datos_dirichlet$Needlesrmse_m3)

datos_dirichlet$totalrmse_m3<-
with(datos_dirichlet,(Stermrmse_m3+Branches_1rmse_m3+Branches_2rmse_m3+Needlesrmse_m
3))

sum(datos_dirichlet$totalrmse_m3)

#### For model 5 -> Model 5: DirichReg(formula = AL ~ I(DBH^2) + Ht_m, data = datos_dirichlet)

```

YOU SHOULD CHANGE FOR MODEL 6,10, AND 15 from this point to the end. Repeat the previous stage but changing to the correct model

```
summary(model5)

fitted(model5)

predict(model5,newdata=data.frame("DBH"=seq(10,10,1000), "Ht_m"=seq(10,10,100)))

residuals(model5,type=c('standardized'))

confint(model5)

confint(model5, exp = TRUE)

logLik(model5)

round(vcov(model5), 3)

#Compute predicted(fitted) percentages of component biomass
datos_dirichlet$Stempp_m5 <- (fitted(model5)[,"Stemp"])
datos_dirichlet$Branches_1pp_m5 <- (fitted(model5)[,"Branches_1_1p"])
datos_dirichlet$Branches_2pp_m5 <- (fitted(model5)[,"Branches_2_1p"])
datos_dirichlet$Needlespp_m5 <- (fitted(model5)[,"Needlesp"])

#Compute predicted(fitted) of component biomass
datos_dirichlet$pStem_m5 <- with(datos_dirichlet,(Stempp_m5 *Totalbiomass))
datos_dirichlet$pBranches_1_m5 <-with(datos_dirichlet,(Branches_1pp_m5 *Totalbiomass))
datos_dirichlet$pbranches_2_m5 <-with(datos_dirichlet,(Branches_2pp_m5 *Totalbiomass))
datos_dirichlet$pNeedles_m5 <-with(datos_dirichlet,(Needlespp_m5 *Totalbiomass))
datos_dirichlet$pbiomastotal_m5 <-with(datos_dirichlet,(pStem_m5 +pBranches_1_m5
+pbranches_2_m5 +pNeedles_m5 ))

# Compute mean bias of component biomass

# -<caution 16 is the number of observations (n)

datos_dirichlet$Stembias_m5=with(datos_dirichlet,(((100/16)*(Stem-pStem_m5)/Stem))

mean(datos_dirichlet$Stembias_m5)

datos_dirichlet$Branches_1bias_m5=with(datos_dirichlet,(((100/16)*(Branches_1-
pBranches_1_m5)/Branches_1))

mean(datos_dirichlet$Branches_1bias_m5)
```



```

datos_dirichlet$Branches_2bias_m5=with(datos_dirichlet,((100/16)*(Branches_2-
pbranches_2_m5)/Branches_2))

mean(datos_dirichlet$Branches_2bias_m5)

datos_dirichlet$Needlesbias_m5=with(datos_dirichlet,((100/16)*(Needles-pNeedles_m5)/Needles))

mean(datos_dirichlet$Needlesbias_m5)

datos_dirichlet$totalbias_m5<-
with(datos_dirichlet,(Stembias_m5+Branches_1bias_m5+Branches_2bias_m5+Needlesbias_m5))

mean(datos_dirichlet$totalbias_m5)

#Compute mean RMSE of component biomass

datos_dirichlet$Stemrmse_m5<-with(datos_dirichlet,sqrt((Stem-pStem_m5)^2)/16)

sum(datos_dirichlet$Stemrmse_m5)

datos_dirichlet$Branches_1rmse_m5<-with(datos_dirichlet,sqrt((Branches_1-
pBranches_1_m5)^2)/16)

sum(datos_dirichlet$Branches_1rmse_m5)

datos_dirichlet$Branches_2rmse_m5<-with(datos_dirichlet,sqrt((Branches_2-
pbranches_2_m5)^2)/16)

sum(datos_dirichlet$Branches_2rmse_m5)

datos_dirichlet$Needlesrmse_m5<-with(datos_dirichlet,sqrt((Needles-pNeedles_m5)^2)/16)

sum(datos_dirichlet$Needlesrmse_m5)

datos_dirichlet$totalrmse_m5<-
with(datos_dirichlet,(Stemrmse_m5+Branches_1rmse_m5+Branches_2rmse_m5+Needlesrmse_m
5))

sum(datos_dirichlet$totalrmse_m5)

### For model 6 -> Model 6: DirichReg(formula = AL ~ I(DBH^2) + I(Ht_m^2), data =
datos_dirichlet)

## This model didnot converge in my script

summary(model6)

fitted(model6)

predict(model6,newdata=data.frame("DBH"=seq(10,10,1000), "Ht_m"=seq(10,10,100)))

residuals(model6,type=c('standardized'))

confint(model6)

```

```

confint(model6, exp = TRUE)

logLik(model6)

round(vcov(model6), 3)

#Compute predicted(fitted) percentages of component biomass
datos_dirichlet$Stempp_m6 <- (fitted(model6)[,"Stemp"])
datos_dirichlet$Branches_1pp_m6 <- (fitted(model6)[,"Branches_1_1p"])
datos_dirichlet$Branches_2pp_m6 <- (fitted(model6)[,"Branches_2_1p"])
datos_dirichlet$Needlespp_m6 <- (fitted(model6)[,"Needlesp"])

#Compute predicted(fitted) of component biomass
datos_dirichlet$pStem_m6 <- with(datos_dirichlet,(Stempp_m6 *Totalbiomass))
datos_dirichlet$pBranches_1_m6 <-with(datos_dirichlet,(Branches_1pp_m6 *Totalbiomass))
datos_dirichlet$pbranches_2_m6 <-with(datos_dirichlet,(Branches_2pp_m6 *Totalbiomass))
datos_dirichlet$pNeedles_m6 <-with(datos_dirichlet,(Needlespp_m6 *Totalbiomass))
datos_dirichlet$pbiomastotal_m6 <-with(datos_dirichlet,(pStem_m6 +pBranches_1_m6
+pbranches_2_m6 +pNeedles_m6 ))

# Compute mean bias of componenet biomass

#caution 16 is the number of observations (n)
datos_dirichlet$Stembias_m6=with(datos_dirichlet,((100/16)*(Stem-pStem_m6)/Stem))
mean(datos_dirichlet$Stembias_m6)

datos_dirichlet$Branches_1bias_m6=with(datos_dirichlet,((100/16)*(Branches_1-
pBranches_1_m6)/Branches_1))
mean(datos_dirichlet$Branches_1bias_m6)

datos_dirichlet$Branches_2bias_m6=with(datos_dirichlet,((100/16)*(Branches_2-
pbranches_2_m6)/Branches_2))
mean(datos_dirichlet$Branches_2bias_m6)

datos_dirichlet$Needlesbias_m6=with(datos_dirichlet,((100/16)*(Needles-pNeedles_m6)/Needles))
mean(datos_dirichlet$Needlesbias_m6)

datos_dirichlet$totalbias_m6<-
with(datos_dirichlet,(Stembias_m6+Branches_1bias_m6+Branches_2bias_m6+Needlesbias_m6))

```

```

mean(datos_dirichlet$totalbias_m6)

#Compute mean RMSE of component biomass
datos_dirichlet$Stemrmse_m6<-with(datos_dirichlet,sqrt((Stem-pStem_m6)^2)/16)
sum(datos_dirichlet$Stemrmse_m6)

datos_dirichlet$Branches_1rmse_m6<-with(datos_dirichlet,sqrt((Branches_1-
pBranches_1_m6)^2)/16)
sum(datos_dirichlet$Branches_1rmse_m6)

datos_dirichlet$Branches_2rmse_m6<-with(datos_dirichlet,sqrt((Branches_2-
pbranches_2_m6)^2)/16)
sum(datos_dirichlet$Branches_2rmse_m6)

datos_dirichlet$Needlesrmse_m6<-with(datos_dirichlet,sqrt((Needles-pNeedles_m6)^2)/16)
sum(datos_dirichlet$Needlesrmse_m6)

datos_dirichlet$totalrmse_m6<-
with(datos_dirichlet,(Stemrmse_m6+Branches_1rmse_m6+Branches_2rmse_m6+Needlesrmse_m
6))
sum(datos_dirichlet$totalrmse_m6)

#### For model 10
summary(model10)
fitted(model10)
predict(model10,newdata=data.frame("DBH"=seq(10,10,1000), "Ht_m"=seq(10,10,100)))
residuals(model10,type=c('standardized'))
confint(model10)
confint(model10, exp = TRUE)
logLik(model10)
round(vcov(model10), 3)

#Compute predicted(fitted) percentages of component biomass
datos_dirichlet$Stempp_m10 <- (fitted(model10)[,"Stemp"])
datos_dirichlet$Branches_1pp_m10 <- (fitted(model10)[,"Branches_1_1p"])
datos_dirichlet$Branches_2pp_m10 <- (fitted(model10)[,"Branches_2_1p"])

```

```

datos_dirichlet$Needlespp_m10 <- (fitted(model10)[,"Needlesp"])
#Compute predicted(fitted) of component biomass
datos_dirichlet$pStem_m10 <- with(datos_dirichlet,(Stempp_m10 *Totalbiomass))
datos_dirichlet$pBranches_1_m10 <-with(datos_dirichlet,(Branches_1pp_m10 *Totalbiomass))
datos_dirichlet$pbranches_2_m10 <-with(datos_dirichlet,(Branches_2pp_m10 *Totalbiomass))
datos_dirichlet$pNeedles_m10 <-with(datos_dirichlet,(Needlespp_m10 *Totalbiomass))
datos_dirichlet$pbiomastotal_m10 <-with(datos_dirichlet,(pStem_m10 +pBranches_1_m10
+pbranches_2_m10 +pNeedles_m10 ))
# Compute mean bias of component biomass
# -<caution 16 is the number of observations (n)
datos_dirichlet$Stembias_m10=with(datos_dirichlet,((100/16)*(Stem-pStem_m10)/Stem))
mean(datos_dirichlet$Stembias_m10)
datos_dirichlet$Branches_1bias_m10=with(datos_dirichlet,((100/16)*(Branches_1-
pBranches_1_m10)/Branches_1))
mean(datos_dirichlet$Branches_1bias_m10)
datos_dirichlet$Branches_2bias_m10=with(datos_dirichlet,((100/16)*(Branches_2-
pbranches_2_m10)/Branches_2))
mean(datos_dirichlet$Branches_2bias_m10)
datos_dirichlet$Needlesbias_m10=with(datos_dirichlet,((100/16)*(Needles-
pNeedles_m10)/Needles))
mean(datos_dirichlet$Needlesbias_m10)
datos_dirichlet$totalbias_m10<-
with(datos_dirichlet,(Stembias_m10+Branches_1bias_m10+Branches_2bias_m10+Needlesbias_m
10))
mean(datos_dirichlet$totalbias_m10)
#Compute mean RMSE of component biomass
datos_dirichlet$Stemrmse_m10<-with(datos_dirichlet,sqrt((Stem-pStem_m10)^2)/16)
sum(datos_dirichlet$Stemrmse_m10)
datos_dirichlet$Branches_1rmse_m10<-with(datos_dirichlet,sqrt((Branches_1-
pBranches_1_m10)^2)/16)
sum(datos_dirichlet$Branches_1rmse_m10)

```

```

dato_dirichlet$Branches_2rmse_m10<-with(datos_dirichlet,sqrt((Branches_2-
pbranches_2_m10)^2)/16)

sum(datos_dirichlet$Branches_2rmse)

datos_dirichlet$Needlesrmse_m10<-with(datos_dirichlet,sqrt((Needles-pNeedles_m10)^2)/16)

sum(datos_dirichlet$Needlesrmse_m10)

datos_dirichlet$totalrmse_m10<-
with(datos_dirichlet,(Stemrmse_m10+Branches_1rmse_m10+Branches_2rmse_m10+Needlesrmse
_m10))

sum(datos_dirichlet$totalrmse_m10)

### For model 15

fitted(model15)

predict(model15,newdata=data.frame("DBH"=seq(10,10,1000), "Ht_m"=seq(10,10,100)))

residuals(model15,type=c('standardized'))

confint(model15)

confint(model15, exp = TRUE)

logLik(model15)

round(vcov(model15), 3)

#Compute predicted(fitted) percentages of component biomass

datos_dirichlet$Stempp_m15 <- (fitted(model15)[,"Stemp"])

datos_dirichlet$Branches_1pp_m15 <- (fitted(model15)[,"Branches_1_1p"])

datos_dirichlet$Branches_2pp_m15 <- (fitted(model15)[,"Branches_2_1p"])

datos_dirichlet$Needlespp_m15 <- (fitted(model15)[,"Needlesp"])

#Compute predicted(fitted) of component biomass

datos_dirichlet$pStem_m15 <- with(datos_dirichlet,(Stempp_m15 *Totalbiomass))

datos_dirichlet$pBranches_1_m15 <-with(datos_dirichlet,(Branches_1pp_m15 *Totalbiomass))

datos_dirichlet$pbranches_2_m15 <-with(datos_dirichlet,(Branches_2pp_m15 *Totalbiomass))

datos_dirichlet$pNeedles_m15 <-with(datos_dirichlet,(Needlespp_m15 *Totalbiomass))

datos_dirichlet$pbiomastotal_m15 <-with(datos_dirichlet,(pStem_m15 +pBranches_1_m15
+pbranches_2_m15 +pNeedles_m15 ))

```

```

# Compute mean bias of component biomass

# -<caution 16 is the number of observations (n)

datos_dirichlet$Stembias_m15=with(datos_dirichlet,((100/16)*(Stem-pStem_m15)/Stem))

mean(datos_dirichlet$Stembias_m15)

datos_dirichlet$Branches_1bias_m15=with(datos_dirichlet,((100/16)*(Branches_1-
pBranches_1_m15)/Branches_1))

mean(datos_dirichlet$Branches_1bias_m15)

datos_dirichlet$Branches_2bias_m15=with(datos_dirichlet,((100/16)*(Branches_2-
pbranches_2_m15)/Branches_2))

mean(datos_dirichlet$Branches_2bias_m15)

datos_dirichlet$Needlesbias_m15=with(datos_dirichlet,((100/16)*(Needles-
pNeedles_m15)/Needles))

mean(datos_dirichlet$Needlesbias_m15)

datos_dirichlet$totalbias_m15<-
with(datos_dirichlet,(Stembias_m15+Branches_1bias_m15+Branches_2bias_m15+Needlesbias_m
15))

mean(datos_dirichlet$totalbias_m15)

#Compute mean RMSE of component biomass

datos_dirichlet$Stermrmse_m15<-with(datos_dirichlet,sqrt((Stem-pStem_m15)^2)/16)

sum(datos_dirichlet$Stermrmse_m15)

datos_dirichlet$Branches_1rmse_m15<-with(datos_dirichlet,sqrt((Branches_1-
pBranches_1_m15)^2)/16)

sum(datos_dirichlet$Branches_1rmse_m15)

datos_dirichlet$Branches_2rmse_m15<-with(datos_dirichlet,sqrt((Branches_2-
pbranches_2_m15)^2)/16)

sum(datos_dirichlet$Branches_2rmse_m15)

datos_dirichlet$Needlesrmse_m15<-with(datos_dirichlet,sqrt((Needles-pNeedles_m15)^2)/16)

sum(datos_dirichlet$Needlesrmse_m15)

datos_dirichlet$totalrmse_m15<-
with(datos_dirichlet,(Stermrmse_m16+Branches_1rmse_m15+Branches_2rmse_m15+Needlesrmse
_m15))

```

```

sum(datos_dirichlet$totalrmse_m15)

##### GRAPHICS #####

## First of all

##MODEL 6 (Repeat for all the model selected in order to obtain the graphics or for the best
model selected change m3 by mX X is the number of the model selected)

datos_dirichlet$rStem_m6<-with(datos_dirichlet,(Stem-pStem_m6))

datos_dirichlet$rBranches_1r_m6<-with(datos_dirichlet,(Branches_1-pBranches_1_m6))

datos_dirichlet$rBranches_2r_m6<-with(datos_dirichlet,(Branches_2-pbranches_2_m6))

datos_dirichlet$rNeedles_m6<-with(datos_dirichlet,(Needles-pNeedles_m6))

datos_dirichlet$rbiomastotal_m6<-with(datos_dirichlet,(Totalbiomass-pbiomastotal_m6))

#save data

save(datos_dirichlet, file="datos_dirichlet.RData")

write.csv2(datos_dirichlet,
file="C:/MEDfOR_Spain/MScThesis_MEDFOR_UVa/Data/First_Work_INIA/Eric
Cudjoe_Palencia.csv")

save.image("datos_dirichlet.RData")

str (datos_dirichlet)

###Graph 1

## HEre the example for m3 model (you should change for other models from here to the end
chagin for the right model number and statistic

## for example changing in 'rStem_mX' changin X for the number of the model selected)

par(mfrow=c(1, 4))

#Loaded library for graph

library(car)

scatterplot(datos_dirichlet$Stem~datos_dirichlet$rStem_m6,col="Black",main="Stem Biomass",
            ylab="Residual (kg)", xlab="Observation (kg)")

scatterplot(datos_dirichlet$Branches_1~datos_dirichlet$rBranches_1r_m6,col="Black",main="Thin
Branch Biomass",
            ylab="Residual (kg)", xlab="Observation (kg)")

```

```
scatterplot(datos_dirichlet$Branches_2~datos_dirichlet$rBranches_2r_m6,col="Black",main="Thick
Branch Biomass",
```

```
ylab="Residual (kg)", xlab="Observation (kg)")
```

```
scatterplot(datos_dirichlet$Needles~datos_dirichlet$rNeedles_m6,col="Black",main="Needle
Biomass",
```

```
ylab="Residual (kg)", xlab="Observation (kg)")
```

```
scatterplot(datos_dirichlet$Totalbiomass~datos_dirichlet$rbiomastotal_m6,col="Black",main="Total
Biomass",
```

```
ylab="Residual (kg)", xlab="Observation (kg)")
```

```
##Graph 2
```

```
## to visualized RMSE versus DBH class
```

```
library(lattice)
```

```
par(mfrow=c(1, 2))
```

```
par(mar=c(1,1,1,1))
```

```
plot(datos_dirichlet$DBH,datos_dirichlet$Stemrmse_m6,ylim=c(-2,2),col=c("blue"), main="Stem
Biomass",
```

```
ylab="RMSE(kg)", xlab="Diameter(cm)")
```

```
plot(datos_dirichlet$DBH,datos_dirichlet$Branches_1rmse_m6,ylim=c(-2,2),col=c("blue"),
main="Thin Branch Biomass",
```

```
ylab="RMSE(kg)", xlab="Diameter(cm)")
```

```
plot(datos_dirichlet$DBH,datos_dirichlet$Branches_2rmse_m6,ylim=c(-2,2),col=c("blue"),
main="Thick Branch Biomass",
```

```
ylab="RMSE(kg)", xlab="Diameter(cm)")
```

```
plot(datos_dirichlet$DBH,datos_dirichlet$Needlesrmse_m6,ylim=c(-2,2),col=c("blue"),
main="Needle Biomass",
```

```
ylab="RMSE(kg)", xlab="Diameter(cm)")
```

```
plot(datos_dirichlet$DBH,datos_dirichlet$totalrmse_m6,ylim=c(-2,2),col=c("blue"), main="Total
Biomass", ylab="RMSE(kg)", xlab="Diameter(cm)")
```

```
#Graph 3
```

```
## to visualized Bias(%) versus DBH class
```



```

plot(datos_dirichlet$DBH,datos_dirichlet$Stembias_m6,ylim=c(-20,20),col=c("green"),main="Stem
Biomass",

      ylab="Bias(%)", xlab="Diameter(cm)")

plot(datos_dirichlet$DBH,datos_dirichlet$Branches_1bias_m6,ylim=c(-
20,20),col=c("green"),main="Thin Branch Biomass",

      ylab="Bias(%)", xlab="Diameter(cm)")

plot(datos_dirichlet$DBH,datos_dirichlet$Branches_2bias_m6,ylim=c(-
20,20),col=c("green"),main="Thick Branch Biomass",

      ylab="Bias(%)", xlab="Diameter(cm)")

plot(datos_dirichlet$DBH,datos_dirichlet$Needlesbias_m6,ylim=c(-
20,20),col=c("green"),main="Needle Biomass",

      ylab="Bias(%)", xlab="Diameter(cm)")

plot(datos_dirichlet$DBH,datos_dirichlet$totalbias_m6,ylim=c(-20,20),col=c("green"),main="Total
Biomass",

      ylab="Bias(%)", xlab="Diameter(cm)")

```

3.2. ANCOVA ANALYSIS FOR SCOTS PINE

BIOMASS ALLOMETRY OF MONOSPECIFIC AND MIXED STANDS

Select the working directory

```
install.packages("tidyverse")
```

```
library(tidyverse)
```

```
library(datasets)
```

```
install.packages("ggplot2")
```

```
library(ggplot2)
```

```
setwd("C:/MEDfOR_Spain/MScThesis_MEDFOR_UVa/Data/Final-Work/ANCOVA")
```

```
data<-read.csv('data.csv', sep=',',dec='.',header=T, na.string='NA')
```

Relationship between independent variables

H-DBH

```
data$logHt_m<- with(data,log(Ht_m))
```

```
data$logDBH<- with(data,log(DBH))
```

```

ggplot(data,aes(x=log(DBH),y=log(Ht_m),col=Variables))+geom_point()+geom_smooth(method="gam")

##### Height-Crown base height

data$logHt_m<- with(data,log(Ht_m))

data$logCrown_base_height<- with(data,log(Crown_base_height))

ggplot(data,aes(x=log(Crown_base_height),y=log(Ht_m),col=Variables))+geom_point()+geom_smooth(method="gam")

##### Height-Crown length

data$logHt_m<- with(data,log(Ht_m))

data$logCrown_length<- with(data,log(Crown_length))

ggplot(data,aes(x=log(Crown_length),y=log(Ht_m),col=Variables))+geom_point()+geom_smooth(method="gam")

#Done.. In conclusion, there is no difference between monospecific and mixed stands.

##### ANALYSIS OF COVARIANCE (ANCOVA)
#####

# Select the working directory

setwd("C:/MEDfOR_Spain/MScThesis_MEDFOR_UVa/Data/Final-Work/ANCOVA")

data<-read.csv('data.csv', sep=',',dec='.',header=T, na.string='NA')

# Pine biomass data: total biomass, stem biomass, branch biomass, needle biomass

library('ggplot2')

library('agricolae')

library(car)

library(multcomp)

library(effects)

library(lattice)

library(colorspace)

data$logDBH<- with(data,log(DBH))

data$logTotalbiomass<- with(data,log(Totalbiomass))

data$logStem<- with(data,log(Stem))

```

```

data$logBranches_2<- with(data,log(Branches_2))

data$logBranches_1<- with(data,log(Branches_1))

data$logNeedles<- with(data,log(Needles))

# Some graphics for visualitation tendencies and developemt versus DBH

ggplot(data,aes(x=log(DBH),y=log(Totalbiomass),col=Variables))+geom_point()+geom_smooth(met
hod="gam")

# Although the graphics shows that could not be any differences between pure & mixed stands
for the same species, we could check if

## it is true for all the biomass compartments (and check it in an ANCOVA analysis)

ggplot(data,aes(x=log(DBH),y=log(Stem),col=Variables))+geom_point()+geom_smooth(method="g
am")

ggplot(data,aes(x=log(DBH),y=log(Branches_2),col=Variables))+geom_point()+geom_smooth(meth
od="gam")

ggplot(data,aes(x=log(DBH),y=log(Branches_1),col=Variables))+geom_point()+geom_smooth(meth
od="gam")

ggplot(data,aes(x=log(DBH),y=log(Needles),col=Variables))+geom_point()+geom_smooth(method=
"gam")

##### ANCOVA analysis for PINE #####

setwd("C:/MEDfOR_Spain/MScThesis_MEDFOR_UVa/Data/Final-Work/ANCOVA")

data<-read.csv('data.csv', sep=',',dec='.',header=T, na.string='NA')

PINE<-subset(data,Variables=="Mixed_pine" | Variables=="Pure_pine")

## Total biomass

#ANCOVA, because we have to see if the differences are only due to the advance in the
development of the trees

vinagremodel<-aov(log(Totalbiomass)~log(DBH)+Variables,data=PINE) # The interaction does not
come out significant, there are no different slopes only intercepts

summary.lm(vinagremodel)

summary(vinagremodel)

vinagremodel

```

```

Anova(vinagremodel,type="III")
adjustedMeans<-effect("Variables",vinagremodel,se=TRUE)
summary(adjustedMeans)
postHocs<-glht(vinagremodel,linfct=mcp(Variables="Tukey"))
summary(postHocs)
confint(postHocs)
plot(vinagremodel)
adjustedMeans$se
summary.lm(vinagremodel)

## To test the assumption of homogeneity of regression slopes we execute:
hoRS<-update(vinagremodel, .~. + log(DBH):Variables)
Anova(hoRS, type = "III")

## STEM
#ANCOVA, because we have to see if the differences are only due to the advance in the
development of the trees
vinagremodel<-aov(log(Stem)~log(DBH)+Variables,data=PINE) # The interaction does not come
out significant, there are no different slopes only intercepts
summary.lm(vinagremodel)
summary(vinagremodel)
vinagremodel
Anova(vinagremodel,type="III")
adjustedMeans<-effect("Variables",vinagremodel,se=TRUE)
summary(adjustedMeans)
postHocs<-glht(vinagremodel,linfct=mcp(Variables="Tukey"))
summary(postHocs)
confint(postHocs)
plot(vinagremodel)
adjustedMeans$se
summary.lm(vinagremodel)

```

```

## To test the assumption of homogeneity of regression slopes we execute:

hoRS<-update(vinagremodel, .~. + log(DBH):Variables)

Anova(hoRS, type = "III")

## Branches 2 biomass (Thick branches)

#ANCOVA, porque hay que ver si las diferencias solo son debidas al adelanto en el desarrollo de
los árboles

vinagremodel2<-aov(log(Branches_2)~log(DBH)+Variables,data=PINE)  #la interacción no sale
significativa, no hay diferentes slopes solo intercepts

summary.lm(vinagremodel2)

vinagremodel2

Anova(vinagremodel2,type="III")

adjustedMeans2<-effect("Variables",vinagremodel2,se=TRUE)

summary(adjustedMeans2)

postHocs2<-glht(vinagremodel2,linfct=mcp(Variables="Tukey"))

summary(postHocs2)

confint(postHocs2)

plot(vinagremodel2)

adjustedMeans2$se

summary.lm(vinagremodel2)

## To test the assumption of homogeneity of regression slopes we execute:

hoRS2<-update(vinagremodel2, .~. + log(DBH):Variables)

Anova(hoRS2, type = "III")

## Branches 1 biomass (Thin branches)

#ANCOVA, porque hay que ver si las diferencias solo son debidas al adelanto en el desarrollo de
los árboles

vinagremodel3<-aov(log(Branches_1)~log(DBH)+Variables,data=PINE)  #la interacción no sale
significativa, no hay diferentes slopes solo intercepts

summary.lm(vinagremodel3)

vinagremodel3

```

```

Anova(vinagremodel3,type="III")
adjustedMeans3<-effect("Variables",vinagremodel3,se=TRUE)
summary(adjustedMeans3)
postHocs3<-glht(vinagremodel3,linfct=mcp(Variables="Tukey"))
summary(postHocs3)
confint(postHocs3)
plot(vinagremodel3)
adjustedMeans3$se
summary.lm(vinagremodel3)

## To test the assumption of homogeneity of regression slopes we execute:

hoRS3<-update(vinagremodel3, .~. + log(DBH):Variables)

Anova(hoRS3, type = "III")

## Needle biomass

#ANCOVA, porque hay que ver si las diferencias solo son debidas al adelanto en el desarrollo de
los árboles

vinagremodel4<-aov(log(Needles)~log(DBH)+Variables,data=PINE)    #la interacción no sale
significativa, no hay diferentes slopes solo intercepts

summary.lm(vinagremodel4)

vinagremodel4

Anova(vinagremodel4,type="III")

adjustedMeans4<-effect("Variables",vinagremodel4,se=TRUE)
summary(adjustedMeans4)
postHocs4<-glht(vinagremodel4,linfct=mcp(Variables="Tukey"))
summary(postHocs4)
confint(postHocs4)
plot(vinagremodel4)
adjustedMeans4$se
summary.lm(vinagremodel4)

## To test the assumption of homogeneity of regression slopes we execute:

```

```

hoRS4<-update(vinagremodel4, .~. + log(DBH):Variables)

Anova(hoRS4, type = "III")

##### ANCOVA analysis for OAK #####

setwd("C:/MEDfOR_Spain/MScThesis_MEDFOR_UVa/Data/Final-Work/ANCOVA")

data<-read.csv('data.csv', sep=',',dec='.',header=T, na.string='NA')

OAK<-subset(data,Variables=="Mixed_oak" | Variables=="Pure_oak")

## Total biomass

#ANCOVA, because we have to see if the differences are only due to the advance in the
development of the trees

vinagremodel<-aov(log(Totalbiomass)~log(DBH)+Variables,data=OAK) # The interaction does not
come out significant, there are no different slopes only intercepts

summary.lm(vinagremodel)

summary(vinagremodel)

vinagremodel

Anova(vinagremodel,type="III")

adjustedMeans<-effect("Variables",vinagremodel,se=TRUE)

summary(adjustedMeans)

postHocs<-glht(vinagremodel,linfct=mcp(Variables="Tukey"))

summary(postHocs)

confint(postHocs)

plot(vinagremodel)

adjustedMeans$se

summary.lm(vinagremodel)

## To test the assumption of homogeneity of regression slopes we execute:

hoRS<-update(vinagremodel, .~. + log(DBH):Variables)

Anova(hoRS, type = "III")

## STEM

#ANCOVA, because we have to see if the differences are only due to the advance in the
development of the trees

```

```

inagremodel<-aov(log(Stem)~log(DBH)+Variables,data=OAK) # The interaction does not come out
significant, there are no different slopes only intercepts

summary.lm(vinagremodel)

summary(vinagremodel)

vinagremodel

Anova(vinagremodel,type="III")

adjustedMeans<-effect("Variables",vinagremodel,se=TRUE)

summary(adjustedMeans)

postHocs<-glht(vinagremodel,linfct=mcp(Variables="Tukey"))

summary(postHocs)

confint(postHocs)

plot(vinagremodel)

adjustedMeans$se

summary.lm(vinagremodel)

## To test the assumption of homogeneity of regression slopes we execute:

hoRS<-update(vinagremodel, .~. + log(DBH):Variables)

Anova(hoRS, type = "III")

## Branches 2 biomass (Thick branches)

#ANCOVA, porque hay que ver si las diferencias solo son debidas al adelanto en el desarrollo de
los árboles

vinagremodel2<-aov(log(Branches_2)~log(DBH)+Variables,data=OAK) #la interacción no sale
significativa, no hay diferentes slopes solo intercepts

summary.lm(vinagremodel2)

vinagremodel2

Anova(vinagremodel2,type="III")

adjustedMeans2<-effect("Variables",vinagremodel2,se=TRUE)

summary(adjustedMeans2)

postHocs2<-glht(vinagremodel2,linfct=mcp(Variables="Tukey"))

summary(postHocs2)

```



```

confint(postHocs2)

plot(vinagremodel2)

adjustedMeans2$se

summary.lm(vinagremodel2)

## To test the assumption of homogeneity of regression slopes we execute:

hoRS2<-update(vinagremodel2, .~. + log(DBH):Variables)

Anova(hoRS2, type = "III")

## Branches 1 biomass (Thin branches)


#ANCOVA, porque hay que ver si las diferencias solo son debidas al adelanto en el desarrollo de
los árboles

vinagremodel3<-aov(log(Branches_1)~log(DBH)+Variables,data=OAK) #la interacción no sale
significativa, no hay diferentes slopes solo intercepts

summary.lm(vinagremodel3)

vinagremodel3

Anova(vinagremodel3,type="III")

adjustedMeans3<-effect("Variables",vinagremodel3,se=TRUE)

summary(adjustedMeans3)

postHocs3<-glht(vinagremodel3,linfct=mcp(Variables="Tukey"))

summary(postHocs3)

confint(postHocs3)

plot(vinagremodel3)

adjustedMeans3$se

summary.lm(vinagremodel3)

## To test the assumption of homogeneity of regression slopes we execute:

hoRS3<-update(vinagremodel3, .~. + log(DBH):Variables)

Anova(hoRS3, type = "III")

## Needle biomass

```

#ANCOVA, porque hay que ver si las diferencias solo son debidas al adelanto en el desarrollo de los árboles

```
vinagremodel4<-aov(log(Needles)~log(DBH)+Variables,data=OAK)    #la interacción no sale significativa, no hay diferentes slopes solo intercepts
```

```
summary.lm(vinagremodel4)
```

```
vinagremodel4
```

```
Anova(vinagremodel4,type="III")
```

```
adjustedMeans4<-effect("Variables",vinagremodel4,se=TRUE)
```

```
summary(adjustedMeans4)
```

```
postHocs4<-glht(vinagremodel4,linfct=mcp(Variables="Tukey"))
```

```
summary(postHocs4)
```

```
confint(postHocs4)
```

```
plot(vinagremodel4)
```

```
adjustedMeans4$se
```

```
summary.lm(vinagremodel4)
```

```
## To test the assumption of homogeneity of regression slopes we execute:
```

```
hoRS4<-update(vinagremodel4, .~. + log(DBH):Variables)
```

```
Anova(hoRS4, type = "III")
```