Ciência Florestal
UFSM

ISSN 1980-5098
Open access

Ci. Fl., Santa Maria, v. 32, n. 3, p. 1659-1683, jul./sep. 2022 • https://doi.org/10.5902/1980509840443
Submitted: 09/10/2019 • Approved: 16/12/2020 • Published: 22/09/2022

Nota Técnica

Application of univariate and multivariate statistical analyzes in clonal selection of *Eucalyptus* spp. for charcoal production

Aplicação de análises estatísticas univariada e multivariada na seleção clonal de *Eucalyptus* spp. para a produção de carvão vegetal

Lucas de Freitas Fialho[I]
Angélica de Cássia Oliveira Carneiro[I]
Clarissa Gusmão Figueiró[II]
Letícia Costa Peres[I]
Antônio Policarpo Souza Carneiro[I]
Paula Gabriella Surdi[I]

[I]Universidade Federal de Viçosa, Viçosa, MG, Brazil
[II]Protermo Engenharia Ltda, São Paulo, SP, Brazil

ABSTRACT

The aim of this study was to select superior materials of *Eucalyptus* spp. using univariate and multivariate statistical analyzes. Twenty-five genetic materials from *Eucalyptus* spp. collected in Itamarandiba, Minas Gerais were used. The properties of wood and charcoal were determined for all genetic materials, in addition to the gravimetric yield. Data were submitted to Scott-Knott hierarchical clustering algorithm as univariate analysis. For the multivariate approach, a combination between principal component analysis and hierarchical cluster analysis was used. Both analyzes were efficient in the selection of genetic materials for charcoal production. According to the Scott-Knott test, genetic materials 9 and 21 were the most suitable to produce charcoal. By means of the multivariate analyzes the most indicated were 9, 10 and 21. The Scott-Knott test allowed the visualization of the results of each quality parameter independently. On the other hand, the multivariate tools enabled the observation of the relation between the properties of wood and charcoal.

Keywords: Statistical tools; Scott-Knott; Principal components; Hierarchical cluster; Charcoal quality improvement

RESUMO

O objetivo deste trabalho foi selecionar os materiais superiores de *Eucalyptus* spp. a partir de análises estatísticas univariada e multivariada. Utilizaram-se 25 materiais genéticos de *Eucalyptus* spp., coletados no município de Itamarandiba, Minas Gerais. Foram determinadas as propriedades da madeira e do carvão vegetal para todos os materiais genéticos, além do rendimento gravimétrico. Os dados foram submetidos ao algoritmo de agrupamento hierárquico Scott-Knott, como análise univariada. Para a realização da abordagem multivariada, utilizou-se uma combinação entre as análises de componentes principais e a análise de agrupamentos hierárquicos. Ambas as análises se mostraram eficientes na seleção de materiais genéticos para a produção de carvão vegetal. De acordo com o teste de Scott-Knott, os materiais genéticos 9 e 21 foram os mais indicados para a produção de carvão vegetal. Por meio das análises multivariadas, os mais indicados foram o 9, 10 e 21. O teste de Scott-Knott permitiu a visualização dos resultados de cada parâmetro de qualidade de forma independente, por outro lado, as ferramentas multivariadas possibilitaram a observação das relações entre as propriedades da madeira e do carvão vegetal.

Palavras-chave: Ferramentas estatísticas; Scott-Knott; Componentes principais; Agrupamento hierárquico; Melhoria da qualidade do carvão vegetal

1 INTRODUCTION

Global concerns regarding the increase in anthropogenic emissions have led to the implementation of government policies to support renewable energies. According to IEA/IRENA (2017), more than 60 countries have plans to support this topic. In Brazil, the National Plan on Climate Change (MME, 2007) and the National Energy Plan 2030 (EPE, 2007) establish targets for the reduction of greenhouse gas emissions and the progressive decrease in use of fossil fuels. Parallel to the environmental impacts caused by non-renewable sources of energy, the decline in its reserves, price instability and security issues have reinforced the need for diversification of the global energy matrix (GARCIA *et al*., 2018).

In this context, biomass appears as an interesting alternative for the various sectors where there is a predominance of fossil fuels. Among the advantages of its use, biomass can be considered carbon neutral, since during its growth, CO_2 is removed from the atmosphere through photosynthesis. In addition, it contains low levels of nitrogen and sulfur, which leads to low NO_x and SO_2 emissions (QIAN *et al*., 2011).

Brazil has a vast availability of biomasses, which, in addition to the hydroelectric power generation already existing in the country, can balance the primary energy demands (WELFLE, 2017). Biomass based on wood resources is one of the main sources capable of supplying part of this market. In this context, Brazil stands out due to its area of planted forests, in addition to its volumetric productivity. In 2018, the total area of planted forests was 7.83 million hectares and an average productivity of 36.0 $m^3ha^{-1}year^{-1}$ (IBA, 2019). Eucalyptus is the main genus used in the Brazilian planted forest segment, representing 72.41% of the entire planted area (IBA, 2019), since it has advantages such as rapid growth and adaptability to edaphoclimatic conditions. It is noteworthy that the maximization of volumetric production must always be associated with a quality raw material for a particular use, which is often not considered by different sectors that produce forests.

In the selection of genetic materials to produce charcoal, Tukey test is mainly used as the comparison analysis of the means between the treatments. This test divides the means of treatments into groups that are not completely distinct, which may result, for example, in the presence of many genetic materials in different groups simultaneously, as observed by different researchers (CARNEIRO *et al.*, 2017; PEREIRA *et al.*, 2016; PEREIRA *et al*., 2013). This phenomenon is called overlapping and, in turn, hinders the final interpretation of the results.

As the complexity of the analyzes increases, there is also an increase in the number of variables to be studied, making their interpretation increasingly difficult. Statistical tools are some of the main approaches used to aid in the interpretation of the characteristics of importance for a given observation. Therefore, the choice of the statistical analysis to be used, as the measure of the magnitude of the differences between the observations of the properties, is essential for obtaining reliable results from the study population, since incorrect analyzes or interpretations of the statistics can result in economic losses (BOOS; STEFANSKI, 2011; COLQUHOUN, 2014).

Some of the alternatives to solve this problem are the use of different methods of statistical analysis, such as the hierarchical clustering algorithm Scott-Knott, or the multivariate analyzes. The Scott-Knott test is a univariate analysis that has the advantage of dividing the genetic materials (treatments) into completely distinct groups, thus not presenting the overlapping problem (JELIHOVSCHI *et al.*, 2014). Multivariate analysis, specifically the combined use of main component analysis and cluster analysis, allows, besides the division of genetic materials (treatments) into completely distinct groups, the interpretation of the relation between the properties studied and the reduction of the dimensionality of the data (COSTA *et al.*, 2014).

In the literature, most studies involving univariate or multivariate analyzes for the quality of wood and charcoal are restricted to a small variety of genetic materials and / or studies that evaluate a small range of variables (Castro *et al*., 2013; Protásio *et al*. , 2013; Protásio *et al.*, 2014; Protásio *et al.*, 2015; Moutinho *et al*., 2017). Therefore, the evaluation of a database involving genetic materials of different species and crosses, in addition to a large number of variables, becomes an analysis closer to that of the genetic improvement programs of companies producing charcoal. In order to assist in the evaluation of the genetic variability of eucalypts materials, it was sought, in this work, through the properties of wood and charcoal, to select the superior materials trough univariate and multivariate statistical analysis.

2 MATERIAL AND METHODS

Twenty-five genetic materials of *Eucalyptus* spp. (Table 1) at 87 months of age, cultivated at 3 x 3 meters spacing, from a clonal test in plots of six lines with four plants per line, belonging to Aperam BioEnergia in the municipality of Itamarandiba, Minas Gerais were used. Three trees of medium diameter were selected for each of the 25 clones, totaling 75 trees. Trees with visual and/or border defects were excluded.

Table 1 – Information on the genetic materials of *Eucalyptus* spp.

Identification	Genetic Materials
1	Eucalyptus cloeziana
2 and 3	Hybrid of *Eucalyptus urophylla* and *Eucalyptus* sp.
4, 5, 6, 7 and 10	Hybrid of *Eucalyptus urophylla* and *Eucalyptus* sp.
8	Hybrid of *Eucalyptus grandis* and *Eucalyptus urophylla*
9	Hybrid of *Eucalyptus urophylla* and *Eucalyptus maidenii*
11	Hybrid of *Eucalyptus urophylla, Eucalyptus* sp. and *E. globulus*
12	Hybrid of *Eucalyptus urophylla, Eucalyptus camaldulensis, Eucalyptus grandis* and *Eucalyptus maidenii*
13	Hybrid of *Eucalyptus urophylla, Eucalyptus camaldulensis, Eucalyptus grandis* and *Eucalyptus globulus*
14	Hybrid of *Eucalyptus urophylla, Eucalyptus camaldulensis, Eucalyptus grandis* and *Eucalyptus* sp.
15	Hybrid of *Eucalyptus camaldulensis, Eucalyptus grandis, Eucalyptus urophylla* and *Eucalyptus* sp.
16, 17, 18, 19, 20, 21, 22 and 23	Hybrid of *Eucalyptus urophylla Eucalyptus camaldulensis, Eucalyptus grandis* and *Eucalyptus* sp.
24 and 25	Hybrid of *Eucalyptus urophylla, Eucalyptus pellita* and *Eucalyptus* sp.

Source: Authors (2020)

To calculate the volume, diameter measurements were taken at each meter of the trunk, from the cutting base of each tree to a diameter of 7 cm. The volume of each section was calculated using the formula proposed by Smalian, obtaining the total volume of each tree by adding the volumes of the sections. The total volume of wood produced per hectare was calculated by multiplying the volume of the selected trees by the number of trees per hectare. Finally, the mean annual increment was calculated (RIBEIRO *et al.*, 2017).

Six discs (0%, DBH - diameter at breast height, 25%, 50%, 75% and 100% of the commercial height of the trunk) were cut from each tree. Firstly, the heartwood and sapwood percentages of each disc were measured. The discs referring to the height of the DBH were used to determine the anatomical properties of the wood. Of the other disks, opposite wedges were cut, passing through the pith, which were used to determine the basic density of the wood. The remainder of each disk was sectioned, forming a composite sample for each tree, one part destined to the production of charcoal and the rest used to determine the other properties of each genetic material.

The determination of the heartwood/sapwood ratio (H/S) was carried out according to Pereira *et al.* (2013). To identify the transition region between the heartwood and the sapwood, a magnifying glass was used. Finally, the average values of the heartwood/sapwood ratio, for each genetic material, were calculated from the weighted average of the H/S ratio of the wooden discs of each tree.

The basic density of the wood was determined using the water immersion method described by ABNT NBR 11941 (ABNT, 2003).

For the morphological analysis of fibers, small fragments were cut in the radial direction in the heartwood region of the discs corresponding to the DBH of each tree. The fibers were individualized by conditioning them in a solution of hydrogen peroxide and acetic acid. From the individualized fibers, temporary slides were made to measure the diameter of the lumen and the width of 30 fibers. The wall fraction was determined according to Pereira *et al.* (2013).

To determine the pore diameter and frequency, samples from the heartwood periphery of the tree DBH were used. Anatomical dehydrated cuts were made in an alcoholic series, stained with safranin solution, and the slides fixed with Entelan. 30 pores were measured for each tree and the frequency in 5 micrographs per tree. To perform the measurements, the Axio-Vision 4.3 software was used (PEREIRA *et al.* 2013).

The elementary chemical composition of the wood was carried out according to the methodology described by DIN EN 15104 (2011).

The structural chemical composition of the wood was determined using a composed sample for each tree. The total extractives content was determined according to TAPPI 204 om-88 (TAPPI, 1996), replacing ethanol/benzene with ethanol/toluene. The soluble and insoluble lignin content was determined according to Gomide and Demuner (1986) and Goldshimid (1971), respectively. The total lignin content was obtained by adding the contents of insoluble lignin and soluble lignin.

The ash content was determined according to ABNT NBR 8112 (ABNT, 1986), adapting the platinum crucible for porcelain crucible, and the temperature of 750 °C to 600 °C. The holocelluloses content was calculated by subtracting from 100 the values of total lignin and extractives.

The higher heating value of the wood was determined using an IKA300 adiabatic calorimetric pump, according to the methodology described by ABNT NBR 8633 (1984).

The carbonization of the genetic materials was carried out in a muffle-type electric oven using approximately 300 grams of oven dried wood at 103 ± 2°C until constant mass. A metal container of approximately 0.003 m^3 was used to conduct the process inside the muffle. The heating control was performed manually in increments of 50°C every 30 minutes, corresponding to a heating rate of 1.67°C/min, until the final temperature of 450°C, remaining stabilized in the latter for 60 minutes. At the end of the process, the charcoal yield was determined by gravimetry.

The apparent relative density of charcoal was determined using the hydrostatic method, according to Veiga *et al.* (2018).

Friability was determined according to the methodology proposed by Dias *et al.* (2015) for charcoal. This sample was rotated at 35 rpm for 15 minutes using an electronic friabilometer (MA-791).

The volatile matter and ash contents of the charcoal were determined according to ABNT NBR 8112 (ABNT, 1986). The fixed carbon content was calculated by subtracting the contents of volatile matter and ash from 100. The higher heating value was determined using an IKA300 adiabatic calorimetric pump, according to the methodology described by ABNT NBR 8633 (1984).

The univariate statistical analysis of the data from the different genetic materials was carried out using software R version 3.4.3 (R CORE TEAM, 2017). The Shapiro-Wilk and Bartlett tests were used to analyze normality and homogeneity of variance, respectively. Once the statistical assumptions were satisfactory for all variables, the database was submitted to analysis of variance (F test, $p <0,05$) and when significant differences were established, the Scott-Knott hierarchical clustering algorithm was applied at 5% significance as an exploratory analysis.

For the calculation of the main components, an algorithm of principal component analysis (PCA) was used, with the aid of R version 3.4.3 software (R CORE TEAM, 2017). The PCA was performed based on the correlation matrix between the wood and charcoal properties (Figure 1), which resulted in the formation of new values for each sample point, the scores of each principal component (CP). The eigenvalues are the values that represent the relative contribution of each variable for each CP.

For the selection of the most important properties in the discrimination of the genetic materials, the ones that had the highest eigenvector in the lower eigenvalue PC were removed from the analysis, without any important information being lost, according to procedures of Jolliffe (2002). This process was performed by removing only one property at a time and repeating the process with each withdrawal. At the end of this process, PCs with an eigenvalue greater than 1 were selected, with the CPs having a greater variance than the mean variance, based on the criterion of analysis of the quality of approximation of the correlation matrix (RENCHER, 2002).

The selection of eucalypts genetic materials into homogeneous groups was performed by analysis of agglomerative hierarchical clustering, using the results obtained in the PCA dimensionality reduction. This analysis was performed based on the method for obtaining groups “complete linkage” (RENCHER, 2002). The Mojena method (1977) was used to determine the optimal number of groups in the dendrogram, and k = 1,25 was used as the stop rule in the definition of the number of groups.

3 RESULTS AND DISCUSSION

Scott-Knott analysis indicated the formation of different numbers of groups for each property of wood and charcoal. In general, the properties of wood and charcoal, whose coefficient of variation was classified as medium, presented the formation of a larger number of groups when compared to properties that presented low coefficient of variation (Tables 2a and 2b).

Table 2a – Genetic materials separated into groups dissimilar to each other, for the properties of wood, by the Scott-Knott test at 5% probability

	Parameters	Group I	Group II	Group III	Group IV	Group V	Group VI
Wood	Mean Annual Increment ($m^3.ha^{-1}.year^{-1}$)	7; 9; 18; 23; 21; 6; 19	3; 20; 11; 22; 10; 12	2; 17; 8; 13; 15	24; 14; 15; 4; 25; 1; 5	-	-
	Heartwood/sapwood ratio (%)	6	13; 7; 11; 9; 3; 16	12; 19	21; 4; 24; 25; 03; 5; 17; 23	1; 8; 18	10; 15; 14; 22
	Basic density ($g.cm^{-3}$)	13; 5; 2	1; 11; 9; 21; 24; 19; 15; 25	22; 4; 14; 10; 7; 20; 23; 18; 16; 17; 12	8; 3; 6	-	-
	Wall fraction (%)	2; 21; 16; 14; 5; 23; 20; 15; 22; 18; 24; 19; 4; 9	17; 13; 12; 8; 6; 3; 10; 11; 7; 1; 25	-	-	-	-
	Pore diameter (μm)	25; 14; 8; 6; 12; 5; 18; 4; 3; 19; 16; 22; 23; 11; 7; 15; 2; 13	17; 24; 9; 20; 21; 10	1	-	-	-
	Pore frequency ($pores.mm^{-2}$)	24	1; 4; 16; 21; 20	10; 5; 18; 15; 9; 25; 14; 22; 17; 12	3; 2; 23; 11; 19; 13; 7; 6; 8	-	-
	Carbon content (%)	13; 10; 12; 9; 25; 5; 19; 22; 24; 6; 21; 11; 20 2; 16; 23; 15	17; 14; 18; 7; 8; 4; 1; 3	-	-	-	-
	Hydrogen content (%)	14; 2; 8; 7; 4; 15; 5; 6; 11; 16; 18; 13; 1; 17; 9	19; 23; 12; 22; 24; 20; 10; 21; 25; 3	-	-	-	-
	Nitrogen content (%)	10; 21; 25; 11; 20; 9; 12; 22; 23; 6; 3	3; 1; 17; 19; 15; 5;2; 14; 16; 24; 13; 8; 18; 4; 7	-	-	-	-
	Sulphur content (%)	1; 18; 8; 14; 4; 2; 7; 6; 5; 17; 16; 19; 9	11; 25; 23; 12; 13; 3; 21; 10; 20; 15; 22; 24	-	-	-	-

Continuation ...

Table 2a – Conclusion

Parameters		Group I	Group II	Group III	Group IV	Group V	Group VI
Wood	Oxygen content (%)	3; 1; 4 8; 7	18; 17; 23; 21; 14; 20; 15; 24; 16; 22; 1; 11; 19; 2; 25; 6; 5; 10; 9; 12; 13	-	-	-	-
	Total extractives content (%)	21	10; 19; 24; 20; 25; 9; 22; 16; 1	8; 7; 23; 18; 6; 13; 12; 5	2; 15; 4; 3; 11; 17; 14	-	-
	Total lignin content (%)	8; 6; 12; 5; 20; 17; 4; 14; 1; 18; 3; 13; 7; 25; 16; 1; 22; 2; 21; 15; 10; 23; 11; 9; 24	-	-	-	-	-
	Holocellulosis content (%)	12; 6; 8; 14; 17; 5; 4; 3; 2; 13; 18	7; 1; 20; 15; 11; 16; 22; 25; 19; 23; 9; 10; 21; 24	-	-	-	-
	Ash content (%)	12; 3; 19; 8; 2; 4; 13; 24; 1; 25; 21; 20; 14; 11; 7; 6; 23; 17; 16; 15; 5; 22; 9; 18; 10	-	-	-	-	-
	Higher heating value (MJ.kg^{-1})	10; 25; 24; 22; 20; 17; 14; 19	21; 9; 12; 18; 16; 23; 2; 6; 13; 3; 8; 5; 7; 15; 1; 11; 4	-	-	-	-

Source: Authors (2020)

Table 2b – Genetic materials separated into groups dissimilar to each other, for the properties of charcoal, by the Scott-Knott test at 5% probability

	Parameters	Group I	Group II	Group III	Group IV	Group V	Group VI
Charcoal	Gravimetric yield (%)	9; 20; 10; 21; 25; 22; 13	3; 1; 10; 12; 5; 4; 24; 16; 8; 23; 7	6; 2; 17; 18; 11; 14; 15	-	-	-
	Apparent density (g.cm-3)	21; 9; 11; 1; 10; 13; 4; 15; 25	23; 2; 6; 12; 8; 18; 7; 24; 17; 16; 19; 22; 14; 3; 20; 5	-	-	-	-
	Friability (%)	6; 7; 23; 9; 24; 13; 16; 25; 3; 11; 19; 12; 2; 17; 4	4; 21; 20; 10; 5; 18; 1; 15; 8; 14; 22	-	-	-	-
	Ash content (%)	15; 1; 22; 23; 13; 14	8; 17; 7; 25; 2; 18; 4; 5; 10; 11; 24; 9; 16; 6; 20; 19; 21; 3; 12	-	-	-	-
	Volatile matter content (%)	1; 10; 21; 24; 13; 8; 6; 22; 7; 12; 19; 5; 9; 23; 3; 20; 16; 15; 4; 25; 14; 2; 18; 11; 17	-	-	-	-	-
	Fixed carbon content (%)	17; 11; 18; 2; 14; 25; 15; 4; 23; 16; 20; 9; 3; 5; 19; 7; 12; 22; 8; 6; 13; 24; 21; 10; 1	-	-	-	-	-
	Higher heating value (MJ.kg-1)	18; 23; 2; 9; 22; 17; 24; 14; 16; 13; 19; 8; 15; 6; 11; 3; 25; 5; 21; 4; 10; 20; 7; 12; 1	-	-	-	-	-

Source: Authors (2020)

For the nitrogen, sulfur, ash content of the wood, and the ash content of the charcoal, few groups are formed despite the average coefficient of variation. According to Qian *et al.* (2011), the elementary chemical composition and the ash content of the wood are similar for the same genetic materials of the genus *Eucalyptus*, when cultivated under the same conditions, thus justifying the formation of a few groups for these properties.

For the basic density of the wood, although it presents a coefficient of variation of 4.8%, it is verified the formation of 4 groups for this property. This is due to the low variability between the replicates of the same genetic material, reducing the residues of the analysis, and evidencing the effect of different genetic materials.

Only one group was formed for the lignin contents of the wood and for volatile matter, fixed carbon, and higher heating value of charcoal. The low coefficient of variation of these properties did not show the effect of the genetic materials.

There is a broad participation of genetic material 9 and 21 in the formation of groups of properties with potential values to produce charcoal. The combination of all the properties qualifies them as superior to produce charcoal, since they fit in the group of greater value of mean annual increment and carbon content. Regarding the charcoal properties, they participated in the groups with the highest values of gravimetric yield and apparent relative density. In addition, these genetic materials were in the group with the lowest values for oxygen content and holocellulose content of wood, as well as friability and ash content of charcoal.

In using a univariate approach to selection with multiple genetic materials and properties, some subjectivities may arise on the final interpretation of results, since genetic materials may not fit into potential groups for all study variables, generating doubts to the researcher / executor. Therefore, the principal component analysis (PCA) can be an alternative for the joint study of the variables, in order to allow visualization of the data structure, to find similarity between samples, to detect anomalous samples (outliers) and to reduce the dimensionality of the set of data (REZENDE, 2017).

From the perspective of the selection of eucalypts genetic materials, the first step for a PCA is to reduce the dimensionality of the data set, resulting in a set of variables that retain most of the data variance (VANTUCH *et al.,* 2016). In the discard of variables, for the genetic materials studied, some properties of the wood were removed: pore diameter; wall fraction; carbon content; hydrogen content; nitrogen content, sulfur content; oxygen content; holocelluloses content; ash content and higher heating value, and the following properties of charcoal: fixed carbon content; volatile matter content; ash content and higher heating value.

The properties of the wood: pore diameter, wall fraction, holocelluloses content and higher heating value, and the properties of charcoal: volatile matter content, fixed carbon content and higher heating value, were removed from the analysis for having a higher eigenvector in the principal component of lower eigenvalue, during the principal component analysis. In addition, these variables had a low coefficient of variation, in relation to the others, which corroborated with their withdrawal during PCA (JOLLIFFE, 2002).

The ash contents of the wood and charcoal were not part of the final database for analysis of main components, since, although they had coefficient of variation superior to the other properties, they were presented within the recommended by Carneiro *et al.* (2017a) for the steel use of charcoal. Therefore, such variables would not be adequate selection criteria for this study and may result in a non-representative result for the selection of genetic materials.

At the end of the process of reducing the dimensionality of the database, eight properties were maintained: mean annual increment; heartwood/sapwood ratio; basic density; frequency of pores; extractive content; gravimetric yield of charcoal; apparent relative density of charcoal; friability of charcoal.

The basic density of wood and the gravimetric yield of charcoal presented low coefficient of variation when compared with the other properties. However, the maintenance of both properties during dimensionality reduction was necessary due to its economic importance, since the density is the technological property of the wood most used by the charcoal companies for the selection of genetic material. The charcoal yield is the main property that provides information on the efficiency of the carbonization process (JUIZO *et al.*, 2017).

Table 3 shows the estimates of eigenvalue, variance, and cumulative variance for the principal component analysis. According to the results, eight main components were obtained, which explained 100% of the variability of the data.

Table 3 – Estimation of eigenvalues, variance and cumulative variance associated with the main components

Parameters	Main Components							
	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8
Eigenvalues	2,38	2,12	1,23	0,91	0,57	0,43	0,22	0,14
Variance (%)	29,8	26,45	15,42	11,42	7,10	5,38	2,71	1,72
Cumulative variance (%)	29,8	56,25	71,67	83,09	90,19	95,57	98,28	100

Source: Authors (2020)

In all principal component analysis applications, one must decide how many major components should be used for an efficient explanation of the data. The use of components with low representability of the database can generate some problems regarding the interpretation of the results. The first one is the risk of presenting a result that does not generalize the population and can then be specific to a single sample. And the second of them is the main component to be dominated by a single variable and is therefore not representative of several variables (RENCHER, 2002).

According to the criterion adopted in this study, the principal components with an eigenvalue greater than 1 were selected, thus extracting the first three, which, in a cumulative form, explained 71.67% of the total variability of the data.

According to Dong and Qin (2018), a high correlation value between the variable and the main component is above 0,7 in module. It is observed for the first principal component that the mean annual increment, heartwood/sapwood ratio and friability had negative correlation (Table 4). This indicates that these three properties together were responsible for 29,80% of the data variation explained by component 1 (Table 3).

The eigenvectors of wood and charcoal properties are presented in the Table 4. It is observed that the total extractive content and the gravimetric yield of charcoal presented a positive correlation with the component 2, indicating that these properties were the main responsible for 26,45% of the data variation. However, for component 3 no high correlation was observed, highlighting only the basic density with a value of -0,67.

Table 4 – Eigenvectors of wood and charcoal properties for the selected principal components

Proprieties	**Eigenvectors**		
	PC 1	**PC 2**	**PC 3**
Mean annual increment (MAI)	-0,7	-0,09	0,4
Heartwood/sapwood ratio (HS)	-0,86	-0,14	-0,29
Basic density (BD)	0,16	0,5	-0,67
Frequency of pores (FREP)	0,51	0,57	-0,03
Extrative content (EXT)	-0,21	0,8	0,43
Yield in charcoal (YC)	-0,26	0,77	0,32
Apparently density of charcoal (AD)	-0,29	0,53	-0,45
Friability (FRIA)	-0,82	0,03	-0,24

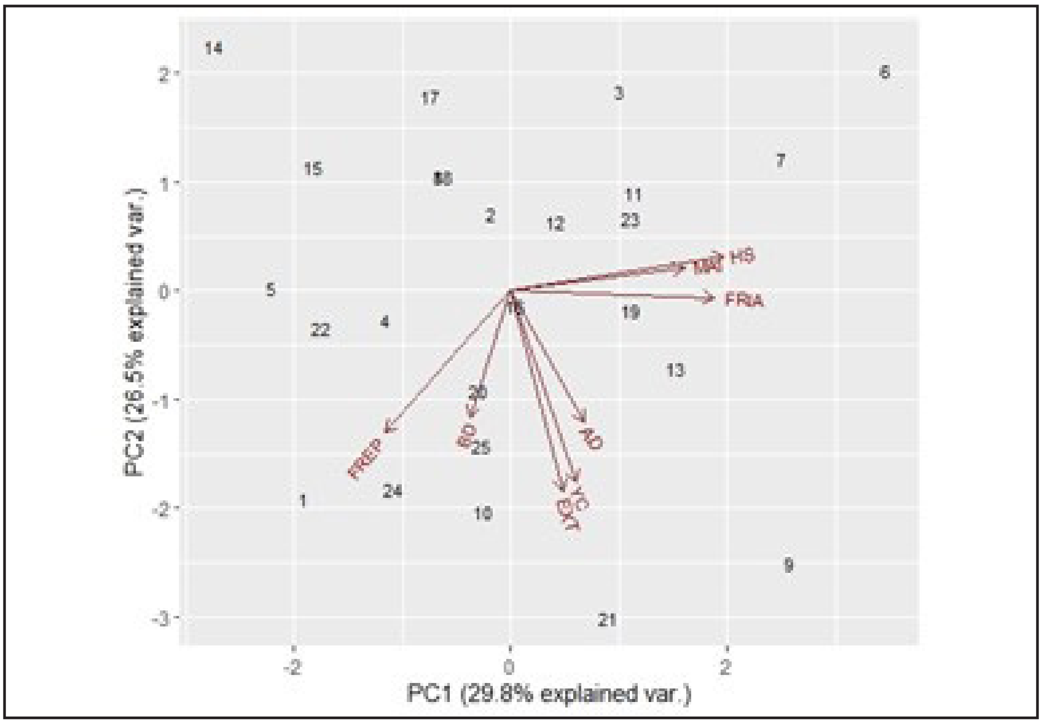
Source: Authors (2020)

In Figures 1 and 2 the genetic materials are represented by dots, identified with their respective numbers, and the variables are represented by their respective vectors. The graphical representation of the biplot type (Figures 1 and 2) allows the visualization of the data matrix, besides making inferences regarding the relation between the variables and existence of similarities between the observations.

The information about the correlation between the study variables can be obtained by means of the cosine of the angle formed between two vectors. If both vectors have the same orientation, there is a positive correlation between the variables. In contrast, if the two vectors have opposing orientations, a negative correlation exists between the variables. If two vectors are almost perpendicular, the correlation is close to zero (JOHNSON; WICHERN, 2007).

It is verified that the correlation between the gravimetric yield and the extractive content was positive, since the angle formed between the vectors is close to 0°, indicating that the presence of these compounds contributed to the increase in the gravimetric yield. Extracts with high carbon content, such as phenolics, may contribute to the increase of the gravimetric yield in charcoal (CARNEIRO *et al.*, 2017b).

Figure 1 – Vectors of the properties of wood and charcoal and dispersion of the scores of the 25 genetic materials in relation to the main components 1 and 2

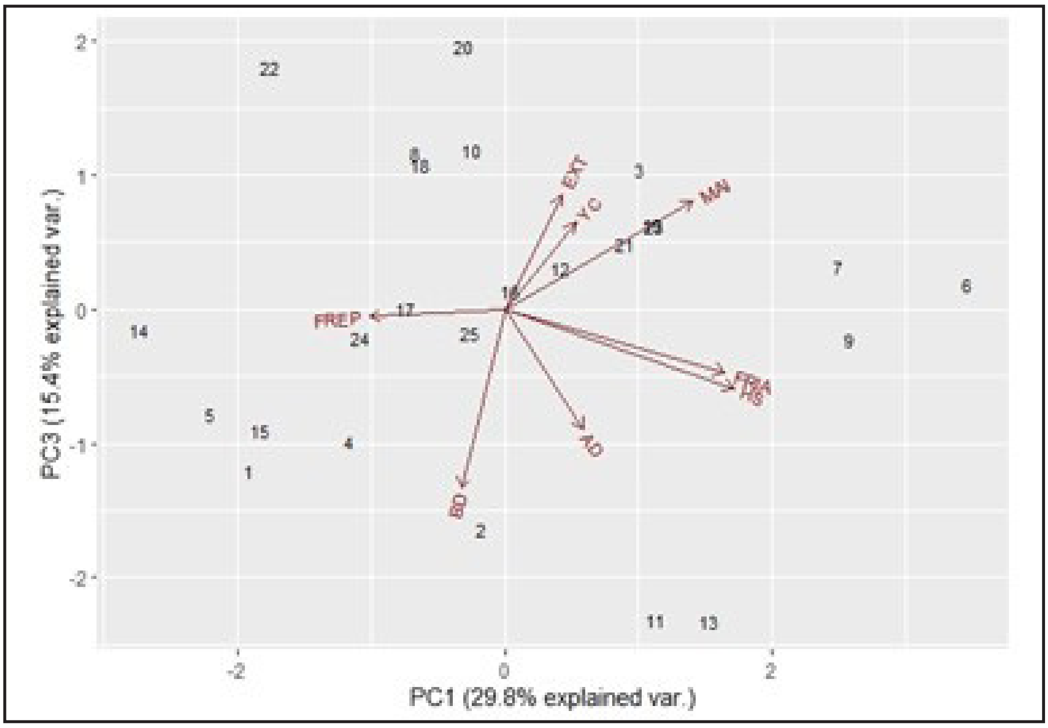


Source: Authors (2020)

In where: MAI = Mean annual increment ($m^3.ha^{-1}.year^{-1}$); HS = Heartwood/sapwood ratio; BD = Basic density ($g.cm^{-3}$); FREP = Frequency of pores ($pores.mm^{-2}$); EXT = Extractive content (%); YC = Yield in charcoal (%); AD = Apparently density of charcoal ($g.cm^{-3}$); FRIA = Friability of charcoal (%).

The vectors formed by the mean annual increment, heartwood/sapwood ratio and friability of charcoal formed, between each other, an angle smaller than 90°, which shows a positive correlation between these properties. A higher heartwood/sapwood ratio makes it difficult to release gases and water vapor from the interior of the wood during carbonization due to the low permeability of the heartwood. This increase in the internal pressure in the wood favors the collapse of its cells, which increases the friability of charcoal (PEREIRA *et al*., 2013). Therefore, it is important that the incremental gain in wood be combined with a breading program that reduces the heartwood/sapwood ratio to improve the quality of the charcoal.

Figure 2 – Vectors of the properties of wood and charcoal and dispersion of the scores of the 25 genetic materials in relation to the main components 1 and 3



Source: Authors (2020)

In where: MAI = Mean annual increment ($m^3.ha^{-1}.year^{-1}$); HS = Heartwood/sapwood ratio; BD = Basic density ($g.cm^{-3}$); FREP = Frequency of pores (pores.mm-2); EXT = Extractive content (%); YC = Yield in charcoal (%); AD = Apparently density of charcoal ($g.cm^{-3}$); FRIA = Friability of charcoal (%).

The friability of charcoal and frequency of wood pores presented a high negative correlation, and the two vectors presented an angle near 180°. A higher pore frequency is associated with a higher permeability of the wood, which may facilitate the release of gases and water vapor from the interior of the wood during carbonization. Thus, lower pressure inside the piece of wood reduces the breakdown of the parenchyma cells, which reduces the friability of the charcoal.

The pore frequency and the basic density of the wood presented a positive correlation, with an angle formed between their vectors less than 90°. According to Baldin *et al.* (2017), a higher pore frequency is associated with the smaller diameter of the pores, which is generally related to a higher wood density.

The basic wood density and the apparent relative density of charcoal presented angles between their vectors less than 90°. The positive correlation between these properties is usually found in most of the studies that correlate wood and charcoal, such as Pereira *et al.* (2013). The positive correlation between the wood basic density and the apparent relative density of charcoal suggests that the selection of a high density wood material could improve the productivity of the charcoal kilns as well as the quality of the charcoal produced, implying lower cost of transportation.

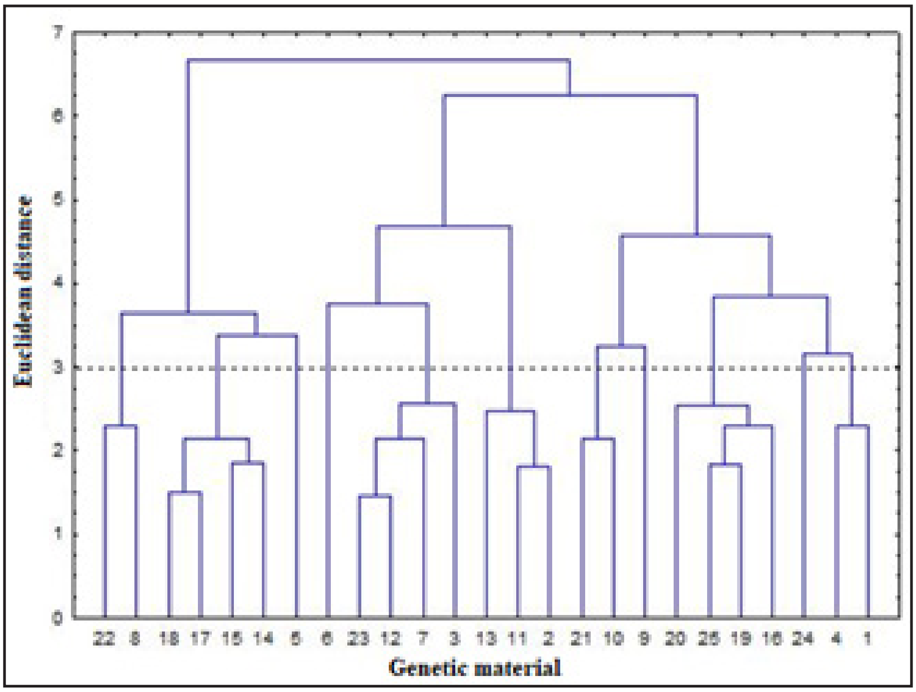
The study of the dissimilarity between the genetic materials by means of the biplot graphical dispersion of the scores generated by the principal components can be realized by the observation of the near points in the graphic, being that the more distant materials are more dissimilar to each other than those nearest (JOHNSON; WICHERN, 2007). Some genetic materials showed greater similarity among themselves, such as materials 21 and 9. It is also possible to observe the formation of some groups that are similar to each other as 18, 17, 15 and 14 (Figures 2 and 3).

The dissimilarity between the genetic materials results, mainly from the properties of wood and charcoal that have higher eigenvectors in the selected main components (JOHNSON; WICHERN, 2007). For the principal component 1 (Table 3), the properties that contributed the most to the dissimilarity of the genetic materials were mean annual increment, heartwood/sapwood ratio and friability; for component 2 the total extractive content and gravimetric yield od charcoal, while for component 3 it was the basic density of the wood.

Cluster analysis is an efficient tool efficient in complementing the results of principal component analysis (VEIGA *et al*., 2018). This analysis is performed with the objective of obtaining patterns in a set of data, through the grouping of the observations. One of the most common approaches to cluster observations is agglomerative hierarchical clusters. In this tool, the analysis begins with “n” groupings, one for each observation, ending with a single group containing all “n” observations.

Figure 3 shows the dendrogram of the genetic materials clusters as a function of the wood and charcoal properties, selected through principal components analysis. Table 6 presents the mean values of the wood and charcoal properties for each group formed by hierarchical group analysis.

Figure 3 – Dendrogram obtained by the “Complete Linkage” method, using the Euclidean distance, obtained through standardized averages of wood and charcoal characteristics for the 25 genetic materials



Source: Authors (2020)

Table 5 – Average values of wood and charcoal properties for the groups formed by the cluster analysis

Group	Genetic material	MAI	HS	BD	FREP	EXT	YC	AD	FRIA
I	22; 8	46,37	0,45	0,55	9,17	5,41	34,91	0,30	5,10
II	18; 17; 15; 14	46,50	0,48	0,55	11,18	3,50	33,03	0,31	5,94
III	5	29,98	0,58	0,60	11,94	4,28	34,75	0,26	6,11
IV	6	55,77	1,37	0,52	7,78	4,62	33,88	0,32	8,84
V	23; 12; 7; 3	54,46	0,82	0,54	9,41	4,41	34,58	0,30	7,81
VI	13; 11; 2	46,01	0,92	0,60	9,07	3,85	34,02	0,34	7,52
VII	21; 10	53,06	0,57	0,57	12,50	7,56	35,93	0,37	6,46
VIII	9	59,96	0,96	0,58	11,25	6,10	37,29	0,37	8,12
IX	20; 25; 19; 16	43,73	0,71	0,56	11,60	6,17	35,35	0,30	7,34
X	24; 4; 1	35,03	0,61	0,57	16,11	5,10	34,74	0,33	6,94

Source: Authors (2020)

In where: MAI = Mean annual increment ($m^3.ha^{-1}.year^{-1}$); HS = Heartwood/sapwood ratio; BD = Basic density ($g.cm^{-3}$); FREP = Frequency of pores ($pores.mm^{-2}$); EXT = Extractive content (%); YC = Yield in charcoal (%); AD = Apparently density of charcoal ($g.cm^{-3}$); FRIA = Friability of charcoal (%).

According to Table 6, the group composed of genetic material 24, 4 and 1 (Group X) and genetic material 5 (Group III) are less suitable to produce charcoal. These groups had a mean annual increment below the Brazilian average, which according to the IBA (2017) was 35.70 $m^3.ha^{-1}.year^{-1}$. It is evidenced that, despite satisfactory results regarding the properties of wood and charcoal, it is necessary that the genetic materials of eucalypts present high volumetric productivity.

Group II, formed by genetic material 18, 17, 15 and 14, and group IV, formed by genetic material 6, although presented adequate values of mean annual increment, basic wood density and apparent density of charcoal, are less indicated because they presented lower gravimetric yield in charcoal, 33.03 and 33.88%, respectively, when compared to the other groups (Table 5). The higher the gravimetric yield, the lower the mass and energy loss of the carbonization process, thus a higher carbon yield implies higher production profits (SANTOS *et al*., 2016).

Groups I, V, VI and IX can be considered as intermediates to produce charcoal (Table 5). In this context, groups I, VI and IX, although showing satisfactory values for most of the properties, genetic materials 22 and 8 (group I) had low relative apparent density, and materials 13, 11 and 2 (group VI) showed a high friability. Group IX, presented low values for the density and friability of charcoal. In relation to the genetic materials 23, 12, 7 and 3 (group V), although they presented high values of mean annual increment, the other properties indicated values of medians to low, emphasizing the low relative apparent density and the high friability of this group of genetic materials.

According to the results of the multivariate analyzes, the most suitable genetic materials to produce charcoal are 21, 10 and 9 (Groups VII and VIII). It is observed that these groups presented high values of mean annual increment, basic wood density, pore frequency and total extractive content. In addition, they presented high values of gravimetric yield and apparent relative density of charcoal, and a low friability of charcoal.

4 CONCLUSION

The Scott-Knott test proved to be efficient for the selection of genetic material to produce charcoal. The genetic material 9 (hybrid of E. *urophylla* and E. *maidenii*) and 21 (hybrid of E. *urophylla*, E. *camaldulensis*, E. *grandis* and *Eucalyptus* sp.) were selected using this statistical tool.

Multivariate analyzes of principal components and clusters were effective for the selection of genetic materials for charcoal production, with materials 9 (hybrid of E. *urophylla* and E. *maidenii*), 10 (hybrid of E. *urophylla* and *Eucalyptus* sp.) and 21 (hybrid of E. *urophylla*, E. *camaldulensis*, E. *grandis* and *Eucalyptus* sp.) are the most indicated from the use of these statistical tools.

The mean annual increment, the heartwood/sapwood ratio, the basic density, the pore frequency, and the extractive content of the wood were the properties that contributed the most to the joint data variance. Among the properties of charcoal, the gravimetric yield, the apparent relative density, and the friability were the most relevant.

REFERENCES

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 8112**: Carvão vegetal - Análise imediata. Rio de Janeiro, 1986. 5 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 8633**: Carvão vegetal - Determinação do poder calorífico. Rio de Janeiro, 1984. 13 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 11941**: Madeira - Determinação da densidade básica. Rio de Janeiro, 2002. 6p.

BALDIN, T.; MARCHIORI, J.N.C.; TALGATTI, M. Anatomia da madeira de *Eucalyptus benthamii* Maiden & Cambage, espécie promissora para o setor de celulose no sul do Brasil. **Balduinia**, Santa Maria, v. 59, p. 10-18, 2017.

BOOS, D.B.; STEFANSKI, L.A. P-Value Precision and Reproducibility. **The American Statistician**, Londres, v. 65, p. 213-221, 2011.

CARNEIRO, A. C. O.; VITAL, B. R.; FREDERICO, P. G. U.; FIALHO, L. F.; FIGUEIRÓ, C. G.; SILVA, C. M. S. Efeito do material genético e do sítio na qualidade do carvão vegetal de madeira de curta rotação. **Floresta**, Curitiba, v. 46, n. 4, p. 473–480, 2017a.

CARNEIRO, A. C. O.; VITAL, B. R.; FREDERICO, P. G. U.; FIGUEIRÓ, C. G.; FIALHO, L. F.; SILVA, C. M. S. Caracterização energética das madeiras de clones de *Eucalyptus* cultivados em diferentes localidades. **Ciência da Madeira**, Pelotas, v. 8, n. 3, p. 127-135, 2017b.

CASTRO, A. F. N. M.; CASTRO, R. V. O.; CARNEIRO, A. de C. O.; LIMA, J. E.; SANTOS, R. C.; PEREIRA, B. L. C.; ALVES, I. C. N. Análise multivariada para seleção de clones de eucalipto destinados à produção de carvão vegetal. **Pesquisa Agropecuária Brasileira**, Brasília, v. 48, n. 6, p. 627-635, 2013.

COLQUHOUN, D. An investigation of the false discovery rate, and the misinterpretation of p-values. **Royal Society Open Science**, v. 19, n. 1, p. 140-216, 2014.

COSTA, T. G.; BIANCHI, M. L.; PROTÁSIO, T. P.; TRUGILHO, P. F.; PEREIRA, A. J. Qualidade da madeira de cinco espécies de ocorrência no cerrado para produção de carvão vegetal. **Cerne**, Lavras, v. 20, n. 1, p. 37-46, 2014.

DIAS, A. F. J.; ANDRADE, C.R.; BRITO, J.O.; MILAN, M. Desdobramento da Função Qualidade (QFD) na Avaliação da Qualidade do Carvão Vegetal Utilizado para Cocção de Alimentos. **Floresta e Ambiente**, Rio de Janeiro, v. 22, n. 2, p. 262-270, 2015.

DONG, Y.; QIN, J. A novel dynamic PCA algorithm for dynamic data modeling and process monitoring. **Journal of Process Control**, v. 68, p. 1-11, 2017.

EMPRESA DE PESQUISAS ENERGÉTICAS (EPE), **Plano Nacional de Energia 2030** (PNE 2030), Rio de Janeiro, 2007.

GARCIA, M.S.; VILPOUX, O.F.; CEREDA, M.P. Distributed electricity generation from sugarcane for agricultural irrigation: A case study from the midwest region of Brazil. **Utilities Policy**, v. 50, p. 207-210, 2018.

GOLDSCHIMID, O. **Ultraviolet spectra**. In: K.V. SARKANEN, C.H. LUDWING. (ed.) Lignins. New York: Wiley Interscience, 1971.

GOMIDE J. L.; DEMUNER B. J. Determinação do teor de lignina em material lenhoso: método Klason modificado. **O Papel**, v.47, p.36-38, 1986.

INDÚSTRIA BRASILEIRA DE ÁRVORES. **Relatório IBÁ 2019: ano base 2018**. 2019, 80p.

International Energy Agency (IEA/IRENA). **Global renewable energy policies & Measures Database**, 2017.

JELIHOVSCHI, E. G.; FARIA, J. C.; ALLAMAN, I. B. 2014. ScottKnott: a package for performing the Scott-Knott clustering algorithm in R. **Tendencias em Matemática Aplicada e Computacional**, v. 15, p. 3–17, 2014.

JOHNSON, R.A.; WICHERN, D.W. **Applied Multivariate Statistical Analysis**. New Jersey: Prentice-Hall, Sixth Edition, 2007.

JOLLIFFE, I.T. Principal component analysis. Ed. 2. New York: Springer Verlag, 2002.

JUIZO, C. G.; LIMA, M. R.; SILVA, D. A. Qualidade da casca e da madeira de nove espécies de Eucalipto para produção de carvão vegetal. **Agrária - Revista Brasileira de Ciências Agrárias**, Recife, v. 12, n. 3, p. 386-390, 2017.

MOJENA, R. Hierarchical grouping methods and stopping rules: an evaluation. **The Computer Journal**. V. 2, p. 359-363, 1977.

NORMEN FÜR HOLZFASERPLATEN SPANPLATTEN SPERRHOLZ. **EN 15104:** Solid biofuels - Determination of total content of carbon, hydrogen and nitrogen - Instrumental methods. Berlin, 2011.

PEREIRA, B. L. C.; CARNEIRO, A. C. O.; CARVALHO, A. M. M. L.; COLODETTE, J. L.; OLIVEIRA, A. C.; FONTES, M. P. F. Influence of Chemical Composition of Eucalyptus Wood on Gravimetric Yield and Charcoal Properties. **BioResources**, v. 8, p. 4574–4592, 2013.

PEREIRA, B. L. C.; CARVALHO, A. M. M. L.; OLIVEIRA, A. C.; SANTOS, L. C.; CARNEIRO, A. C. O.;

MAGALHÃES, M. A. Efeito da carbonização da madeira na estrutura anatômica e densidade do carvão vegetal de *Eucalyptus*. **Ciência Florestal**, Santa Maria, v. 26, n. 2, p. 545-557, 2016.

PROTÁSIO, T. P.; COUTO, A. M.; REIS, A. A.; TRUGILHO, P. F. Seleção de Clones de Eucalyptus para a produção de carvão vegetal e bioenergia por meio de técnicas univariadas e multivariadas. **Scientia Forestalis**, Piracicaba, v. 41, n. 97, p. 15-28, 2013.

PROTÁSIO, T. P.; GOULART, S. L.; NEVES, T. A.; TRUGILHO, P. F.; RAMALHO, F. M. G.; QUEIROZ, L. M. R. S. B. Qualidade da madeira e do carvão vegetal oriundos de floresta plantada em Minas Gerais. **Pesquisa Florestal Brasileira**, Brasília, v. 34, n. 78, p. 111-123, 2014.

PROTÁSIO, T. P.; COUTO, A. M.; REIS, A. A.; TRUGILHO, P. F.; GUIMARÃES, J. B.; LIMA, P. H.; SILVA, M. M. O. Avaliação tecnológica do carvão vegetal da madeira de clones jovens de *Eucalyptus grandis* e *Eucalyptus urophylla*. **Scientia Forestalis**, Piracicaba, v. 41, n. 97, p. 15-28, 2015.

MINISTÉRIO DE MINAS E ENERGIA (MME). Plano Nacional Sobre Mudança Do Clima - Brasil, Brasília, Brasil, 2007.

MOUTINHO, V. H. P.; TOMAZELLO, M.; BRITO, J. O.; BALLARIN, A. W.; ANDRADE, F. W. C.; CARDOSO, C. C. Caracterização e correlação estatística entre as propriedades físicas e Mecânicas do carvão de clones de eucalyptus e corymbia. **Ciência Florestal**, Santa Maria, v. 27, n. 3, p. 1095-1108, 2017.

QIAN, F. P.; CHYANG, C. S.; HUANG, K. S.; TSO, J. Combustion and NO emission of high nitrogen content biomass in a pilot-scale vortexing fluidized bed combustor. **Bioresource Technology**, v. 102, n. 2, p. 1892-1898, 2011.

R Core Team. 2017. R: **A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing. Vienna, Austria. https://www.Rproject.org/.

RENCHER, A. C. **Methods of Multivariate Analysis**. John Wiley & Sons, Inc. Publication, 2ed, 2002.

REZENDE, F. C. Transformações na cientificidade e o ajuste inferencial na Ciência Política: argumento e evidências na produção de alto fator de impacto. **Revista de Sociologia e Política**, Curitiba, v. 25, n. 63, p. 103-138, 2017.

RIBEIRO, M.D.S.B.; JORGE, L.A.B; MISCHAN, M.M.; SANTOS, A.L.; BALLARIN, A.W. Avaliação da produção de biomassa do fuste de um clone híbrido de eucalipto sob diferentes espaçamentos. **Ciência Florestal**, Santa Maria, v. 27, n. 1, p. 31-45, 2017.

SANTOS, R. C.; CARNEIRO, A. C. O.; VITAL, B. R.; CASTRO, R. V. O.; VIDAURRE, G. B.; TRUGILHO, P. F.; CASTRO, A. F. N. M. Influência das propriedades químicas e da relação siringil/guaiacil da madeira de eucalipto na produção de carvão vegetal. **Ciência Florestal**, Santa Maria, v. 26, n. 2, p. 657-669, 2017.

TECHNICAL ASSOCIATION OF THE PULP AND PAPER INDUSTRY. **TAPPI test methods T 204 om-88: solvent extractives of wood and pulp**. In: TAPPI Standard Method. Atlanta, USA, 2001.

VANTUCH, T.; SNASEL, V.; ZELINKA, I. Dimensionality reduction method's comparison based on statistical dependencies. **Procedia Computer Science**, v. 83, p. 1025-1031, 2016.

VEIGA, T. R. L. A.; LIMA, J. T.; MONTEIRO, T. C.; DESSIMONI, A. L. A.; ROCHA, M. F. V. Propriedades mecânicas de amostras individualizadas da madeira e do carvão de Eucalyptus urophylla e de Corymbia citriodora. **Scientia Forestalis**, Piracicaba, v. 46, n. 117, p. 107-114, 2018.

WELFLE, A. Balancing growing global bioenergy resource demands - Brazil's biomass potential and the availability of resource for trade. **Biomass and Bioenergy**, v.105, p.83-95, 2017.

Authorship Contribution

1 – Lucas de Freitas Fialho

Forestry Engineer, Dr., Specialist in Forests

https://orcid.org/0000-0001-8002-9214 • l.freitasfialho@gmail.com

Contribuição: Conceptualization, Methodology, Investigation, Software, Writing – original draft, Writing – review & editing

2 – Angélica de Cássia Oliveira Carneiro

Forestry Engineer, Dr., Professor

https://orcid.org/0000-0002-5992-3059 • cassiacarneiro1@gmail.com

Contribuição: Conceptualization, Methodology, Project administration, Resources, Supervision

3 – Clarissa Gusmão Figueiró

Forestry Engineer, Dr., Operational Excellence Coordinator

https://orcid.org/0000-0002-8572-554X • clarissagfigueiro@gmail.com

Contribuição: Investigation, Software, Writing – original draft, Writing – review & editing

4 – Letícia Costa Peres

Forestry Engineer, MSc.

https://orcid.org/0000-0001-7949-0628 • leticiacostaperes@gmail.com

Contribuição: Investigation, Writing – review & editing

5 – Antônio Policarpo Souza Carneiro

Zootechnist, Dr., Professor

https://orcid.org/0000-0002-9043-3242 • policarpo@ufv.br

Contribuição: Methodology, Software, Validation

6 – Paula Gabriella Surdi

Forestry Engineer, Dr., Researcher

https://orcid.org/0000-0003-0434-9632 • paulasurdi@gmail.com

Contribuição: Validation, Writing – review & editing

How to quote this article

Fialho, L. F.; Carneiro, A. C. O.; Figueiró, C. G.; Peres, L. C.; Carneiro, A. P. S.; Surdi, P. G. Application of univariate and multivariate statistical analyzes in clonal selection of Eucalyptus spp. for charcoal production. Ciência Florestal, Santa Maria, v. 32, n. 3, p. 1659-1683, 2022. DOI 10.5902/1980509840443. Available from: https://doi.org/10.5902/1980509840443.