Industrial Crops & Products 108 (2017) 806–813

Contents lists available at ScienceDirect

Industrial Crops & Products

journal homepage: www.elsevier.com/locate/indcrop

Research Paper

High-performance prediction of macauba fruit biomass for agricultural and industrial purposes using Artificial Neural Networks

Carla Aparecida de O. Castro[a], Rafael T. Resende[a,*], Kacilda N. Kuki[c], Vinícius Q. Carneiro[b], Gustavo E. Marcatti[d], Cosme Damião Cruz[b], Sérgio Y. Motoike[c]

[a] Department of Forestry, Universidade Federal de Viçosa/UFV, Av. Peter Henry Rolfs, s/n 36570-000, Viçosa, MG, Brazil
[b] Department of Statistics, Universidade Federal de Viçosa/UFV, Av. Peter Henry Rolfs, s/n, 36570-000, Viçosa, MG, Brazil
[c] Department of Plant Science, Universidade Federal de Viçosa/UFV, Av. Peter Henry Rolfs, s/n, 36570-000, Viçosa, MG, Brazil
[d] Department of Agricultural Sciences, Universidade Federal de São João Del-Rei/UFSJ, MG424 road, KM 47, n 202, 35701-970, Sete Lagoas, MG, Brazil

ARTICLE INFO

Keywords:
ANN
Prediction method
Dry biomass
Oil contents
Yield
Macaw palm

ABSTRACT

Biomass estimation plays of crucial role in agriculture and agro-based industries. The macauba, *Acrocomia aculeata* (Jacq.) Lood., ex Mart., is a palm species that has been a focal point for research and development of an alternative biomass-bioenergy crop for the tropics. The macauba fruit components (exocarp, mesocarp, endocarp and seed/kernel) present different constitutional characteristics and their biomass determination, by traditional methods, is labor-consuming. Therefore, the validation of procedures that can streamline this process is relevant, since it can reduce costs and time for both breeding programs and industries. This study tested the efficacy of Artificial Neural Networks (ANN) on biomass prediction of the macauba fruit components by comparing it to the multiple linear regression method. The data used came from fruits collected in 18 localities, distributed throughout the state of Minas Gerais, Brazil. According to their provenance, the matrices were clustered into two groups with the k-means method for posterior ANN cross-validation. Each group was interchangeably used for both training and validation purposes. The ANN was more efficient than multivariate linear model in the predictions of dry weight of the fruiťs four components and oil content of the mesocarp and seed. As for variables related to dry weight, ANN reached 98% predictive accuracy (i.e., 98% accuracy of the value predicted by the network), and for variables related to oil contents, accuracy was around 90%. Additionally, non-invasive measurements of the fruit (i.e., low-cost and low-time measurement variables) were adequate enough to predict most of the variables of interest. These results show the ANN's prediction potential, saving time and efforts for the consolidation of macauba as a crop.

1. Introduction

Countries that are signatories of the COP 21 (21st United Nations Conference on Climate Change) face the great challenge of changing their energy matrix, replacing non-renewable sources by renewables ones. This paradigm shift will inevitably increase the demand for biofuels in the world, including biodiesel and bio-kerosene from biomass-bioenergy crops. As a global agreement, biofuel production from plant materials should avoid the deleterious effects of direct and indirect land use change (Lapola et al., 2010; Van der Laan et al., 2016). Therefore, raw materials with high energy density, which adapt to production systems with low environmental impact, are fundamental to sustainable production of biofuels (Lapola et al., 2010). The macauba palm, *Acrocomia aculeata*, fulfills these requirements (Montoya et al., 2016).

Macauba, an oleiferous species of the Tropical America and one of the most conspicuous palm in Brazil, shows great biomass yield potential (Evaristo et al., 2016a). Being a drupaceous species, macauba fruit structure, composition and oil content resemble those of the African oil palm fruit, *Elaeis guineensis* Jacq. (Del Río et al., 2016; Montoya et al., 2016). The macauba mesocarp oil has diverse industrial purposes, including the production of biodiesel, bio-kerosene, oleochemicals and cosmetics (Montoya et al., 2016; Pires et al., 2013). But unlike the African palm, the macauba displays high rusticity to drought, acid soils and fire (Bicalho et al., 2016; Pires et al., 2013); so much so it populates tropical areas with pronounced rainfall seasonality, like the Brazilian savannah ecoregions (Cerrado) (Lanes et al., 2014). Hence, macauba cultivation in degraded areas, such as abandoned pastures, or in integrated systems, such as inter-cropping and silvipastoral systems, is feasible (Motoike and Kuki, 2009). Both are sensible farming practices, with low environmental impact (Lanes et al., 2016). Another

* Corresponding author.
E-mail address: rafael.tassinari@gmail.com (R.T. Resende).

http://dx.doi.org/10.1016/j.indcrop.2017.07.031
Received 29 March 2017; Received in revised form 14 July 2017; Accepted 16 July 2017
0926-6690/

advantage is the likelihood of using all fruit parts – the exocarp (husk), mesocarp (pulp), endocarp (nut) and seed (kernel) – to produce co-products with great economic value and high energy density. The solid residues, after the oil extraction, can provide protein and fiber cakes, charcoal and briquettes (Evaristo et al., 2016c; Pires et al., 2013). For these reasons, macauba is a workable and sustainable non-food crop alternative.

Most of the exploitation of macauba is by means of extractivism in natural populations (Evaristo et al., 2016a; Pires et al., 2013). However, as investments started, commercial plantations have being established in the southeast region of Brazil. To maximize the economic value of the future commercial crops, the development and use of elite macauba plants is required. Cultivars are usually obtained through breeding programs, which often need information on genetic parameters of the species (Lynch and Walsh, 1998). Therefore, the prospect of acquiring estimates of biometric variables of the fruit parts, related to oil content and dry weight, is of relevance for macauba. Biomass predictions would help to set up best breeding strategies. However, the productivity characteristics of fruits – such as *E. guineensis* (Legros et al., 2009) e *Jatropha curcas* L. (Singh et al., 2013) – are controlled by climatic variations and soil properties. In addition, for species with broad distribution, such as macauba, the centers of domestication and origin may not coincide (Lanes et al., 2016). Since Brazil is a country of continental dimensions, with marked climatic gradation, the influence of the environment should be considered when assessing the biomass of macauba fruit.

A fruiťs metrics dataset is useful in evaluating the productivity of commercial plantations and in securing the bases to market their goods (Ciconini et al., 2013). Attaining these values, however, is time and labor demanding. It often needs the individualization of the fruit fractions, for later drying and/or physico-chemical analysis (Coimbra and Jorge, 2011a; Mazzottini-dos-Santos et al., 2015). Thus, analytical strategies that reduce time and costs expenditures, such as indirect prognostics, may increase the efficiency to obtain biometric values of macauba fruits. The approach through Computational Intelligence, using Artificial Neural Networks (ANN) as tools, is an efficient form of prediction. The method reliability is based on many studies that sought to minimize the handling and capture of allometric variables while providing accurate measurements (Baruah et al., 2017; Cheok et al., 2012; Hajar and Vahabzadeh, 2014; Silva et al., 2014). The ANN have a non-linear approach inspired in the human brain operation and structure and, as such, are capable of learning data patterns, including those with interferences and those consisted of incomplete or contradictory events (Silva et al., 2014).

For all that has been said, this study evaluated: (i) the influence of the macauba matrices' provenance on the quantitative traits of the fruit and (ii) the efficiency of ANN for predicting fruit biomass as an alternative to multiple linear regression analysis. The knowledge will help understand the morphological diversity of the species and, equally important, mediate the prognosis of the values referring to the fruit constituents of industrial interest.

2. Materials and methods

2.1. Plant material and provenances

The work used the biometric data of 543 fruits from 172 macauba mother-threes (matrices, 7–25 years old) collected in 18 sites throughout the state of Minas Gerais/Brazil. The fruits were collected between November 2013 and February 2014, directly from the bunches, as soon as natural abscission started. At this stage, the fruits are considered matured but not ripened (Montoya et al., 2016). On average, 9.6 matrices were tested per locality (Table 1). For each matrix, fruit variables were measured in different numbers of repetitions (2–4). Using the cartographic coordinates of the sites of origin, it was possible to obtain information of their meteorological conditions (Table 1), available in WorldClim – global climate records (Hijmans et al., 2005). The average temperature was 21.7 °C and average precipitation was 1306.1 mm^3 throughout the year.

2.2. Evaluation of macauba fruit physico-chemical variables

Eleven variables obtained from the whole fruit and its individual parts were used in the study (Table 2). Based on them, two groups of variables were established:

a *Predictive* – variables that were easily measured with a minimum time spent (between 10 s to 1 min), *viz.* Fruit fresh weight (FFW), Fruit radial diameter (FRD), Fruit axial diameter (FAD), Endocarp width 01 (EW1), Endocarp width 02 (EW2).
b *Predicted* – variables that required more time (4–5 days) and resources to be measured, *viz.* Husk or exocap dry weight (HDW), Pulp or mesocarp dry weight (PDW), Endocarp dry weight (EDW), Kernel or seed dry weight (KDW), Oil content in the kernel (KOC) and Oil content in the pulp (POC).

The predictive variables were classified according to the way in which the set of measurements were obtained: (i) *Non-destructive procedure* – external measures of the fruit, using only three variables from intact fruits i.e., FFW, FAD and FRD; and (ii) *Semi-destructive procedure* – external and internal measures of the fruit, using five variables, three from intact fruits (FFW, FAD and FRD) and two from sectioned fruits (EW1, EW2). These two procedures intended to reveal how much the input variables are relevant to achieve results with better accuracy using multiple regression and ANN. Besides that, it will allow the choice of the most effective strategy for fruit handling.

The non-destructive measurements were carried out as follows: FFW was obtained by individually weighing clean fresh fruits in an electronic scale (0.01 g precision). The FAD and FRD were measured with a digital caliper (0.01 mm precision). For the destructive measurements, fresh fruits were cut in half (radial diameter) with a stainless-steel saw and the EW1 and EW2 randomly assessed using the precision digital caliper (Fig. 1).

To obtain the dry weight of the fruit's parts (HDW, PDW, EDW and KDW), whole fresh fruits were kept in a ventilated kiln at 120 °C for 48 h, as established in previous essays. Afterwards, the fruits were dismantled – using a manual press and a stainless-steel knife – and each fraction conditioned at 65 °C – to allow additional drying but without tissue combustion – and later weighing. The POC and KOC were estimated after extraction of the oils in Soxhlet extractor, using hexane (P.A.) as solvent Montoya et al. (2016).

2.3. Statistical analysis and procedures description

2.3.1. Evaluation of fruit provenance effect over measured variables

The fruits used in this study were collected on sites distributed in a wide range, from north to south of Minas Gerais state (Fig. 2). Therefore, we preliminarily sought to evaluate the effect of the provenance on fruit morphology and, in case of statistical significance, establish groups. In addition, repeatability coefficients (ρ) of fruits observations of a single matrix and the experimental coefficient of variation (*CV* %) were calculated through the adjustment of the mixed-effects model presented in Eq. (1):

$$y = Xb + Zp + \varepsilon, \tag{1}$$

where, y are the observed values of morphological variables of the macauba fruits (i.e., FFW, FRD, FAD, EW1 and EW2); b are fixed effects of localities added to the overall mean; p are the random effects of matrices (or permanent environment effect), with variance structure as $p|\sigma_p^2 \sim N(0, I\sigma_p^2)$; ε is the residual component, with variance structure as $\varepsilon|\sigma_\varepsilon^2 \sim N(0, I\sigma_\varepsilon^2)$; where in I is an identity matrix; X and Z are incidence matrices on fixed and random effects, respectively. The