Computers and Electronics in Agriculture 152 (2018) 401–408

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Bee-inspired RBF network for volume estimation of individual trees

Eugenio Monteiro da Silva Jr.[a,*], Renato Dourado Maia[a], Christian Dias Cabacinha[b]

[a] Computer Science Department, State University of Montes Claros, MG, Brazil
[b] Institute of Agrarian Science, Federal University of Minas Gerais, MG, Brazil

ARTICLE INFO

Keywords:
Planted forest
Eucalyptus
Neural networks
Clustering
cOptBees

ABSTRACT

The *Eucalyptus* is the most cultivated kind of tree in Brazil because it has adapted to the climate and has great importance for the industry. In cultivated forests, the wood volume is essential information to the forest management. Therefore, that information must be estimated as precisely as possible. There are several descriptive mathematical models which were developed for that purpose. However, Computational Intelligence techniques have been used in order to facilitate that process and substitute the volume models. Sundry works have proposed the use of Artificial Neural Networks for wood volume estimation, but there is a type of neural network, the Radial Basis Function – RBF, that can be designed automatically by clustering algorithms. This work presents the application of RBF networks automatically generated by the cOptBees clustering algorithm in the estimation of *Eucalyptus* volume and compares the results to the MLP networks and the classic models at the same dataset. The cOptBees is a clustering algorithm inspired by the behavior of bees which allows the number of clusters to be found automatically. To evaluate the various factors that can influence the quality of the results provided by RBF, the tests consider three training algorithms, three activation functions and three heuristics to define the spread. Besides the RBF generated by cOptBees, were evaluated another two types of RBF: randomly and k-means generated. In the volume estimation, the results indicate that neural networks and classical equations are equivalent to each other when there is high availability of data. However, when there are few training samples, the classical models performed better. Nevertheless, RBF networks are a viable alternative due to its ease of configuration and generalization capability.

1. Introduction

Because it was adapted well to the Brazil's climate and has great applicability as raw material for the industry, the *Eucalyptus* is the most cultivated tree genus in the Brazilian territory (IBA, 2017). In this activity, it is of fundamental importance to know the volume of wood produced.

It is common to select one of several existing equations to provide estimates of volume of wood. Most of these equations have as input the diameter at breast height and the total height of sample trees and have coefficients that are adjusted by means of the rigorous scaling of a certain number of trees. After adjusting some equations, the best equation, according to statistical criteria, is selected and applied to the problem.

Although the traditional estimation method has shown good results in most cases, recent works have demonstrated the interest of forest engineering in Computational Intelligence. Among the Computational Intelligence techniques, Artificial Neural Networks (or Neural Networks) can be perfectly adapted to the described problem. Artificial Neural Networks (ANN) are inspired by biological neural networks and have a massively parallel distributed structure and the capacity of learning with examples and generalizing. Generalization capacity refers to the fact that Neural Networks provide coherent result for input that was never presented in the training phase (Haykin, 1999).

Various works, like (Gorgens et al., 2009; Ozcelik et al., 2010; Binoti et al., 2014; Ozcelik et al., 2014; Sanquetta et al., 2018), use Neural Networks as a method to estimate the volume in cultivated forests. Another example is (Lacerda et al., 2018), that estimate volume in native trees. In the specialized literature the use Neural Networks to volume estimation is not a novelty. However, Computational Intelligence is in constant evolution and recently developed algorithms have not been yet used in this area. The theory suggests that MLP and RBF are equivalent in generalization capacity (Haykin, 1999) and RBF presents some advantages in terms of architecture. Some paper like (Zhang et al., 2016; Wang et al., 2018) explains more about MLP networks. The RBF network has an architecture with fewer parameters than MLP. According to Blanco et al. (2013), this kind of network requires fewer training samples and can be trained faster than MLP's.

* Corresponding author.
E-mail addresses: eugeniomonteiro@ufmg.br (E.M. da Silva), renato.dourado@unimontes.br (R.D. Maia), cabacinha@ica.ufmg.br (C.D. Cabacinha).

https://doi.org/10.1016/j.compag.2018.07.036
Received 1 February 2018; Received in revised form 24 June 2018; Accepted 27 July 2018
Available online 02 August 2018
0168-1699/

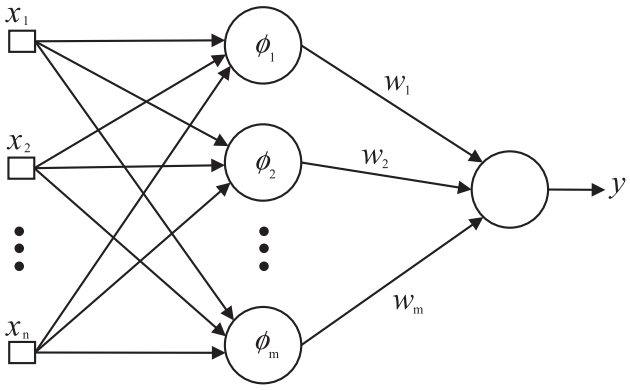

Fig. 1. Typical RBF architecture.

Moreover, the default architecture of RBF can be defined by clustering algorithms removing from the human operator this responsibility.

1.1. RBF networks

Differently to the MLP networks, that can have one or more hidden layers, RBF networks are often defined with only one hidden layer. Fig. 1 presents a typical architecture of a RBF network with n inputs and m neurons.

The neurons of the hidden layer implements a radial basis function. According to Haykin (1999), common functions applied in RBF are: gaussian (Eq. (1)), multiquadric (Eq. (2)) and thin-plate spline (Eq. (3))

$$\phi(u) = e^{-\frac{v^2}{2\sigma^2}} \tag{1}$$

$$\phi(u) = \sqrt{(v^2 + \sigma^2)} \tag{2}$$

$$\phi(u) = v^2 \log v \tag{3}$$

where $v = \|x-\mu\|$, which is usually defined as the Euclidean distance, x is the input vector and μ and σ are, respectively, the center and the spread (or radius) of the radial basis function. Fig. 2 shows two 2D gaussian functions with different spread values (Pazouki et al., 2015).

RBF networks can also be defined by matrix formulation to a better comprehension of how it works. Consider three matrices **G**, **W**, and **D** as, respectively, the result matrix of the hidden layer, the matrix of synaptic weights and the matrix of desired outputs. The goal of the training is find **W**, where:

$$\mathbf{G} \times \mathbf{W} = \mathbf{D} \tag{4}$$

$$\begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,m} \\ \phi_{2,1} & \phi_{2,2} & \cdots & \phi_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n,1} & \phi_{n,2} & \cdots & \phi_{n,m} \end{bmatrix} \times \begin{bmatrix} w_{1,1} \\ w_{2,1} \\ \vdots \\ w_{m,1} \end{bmatrix} = \begin{bmatrix} d_{1,1} \\ d_{2,1} \\ \vdots \\ d_{m,1} \end{bmatrix} \tag{5}$$

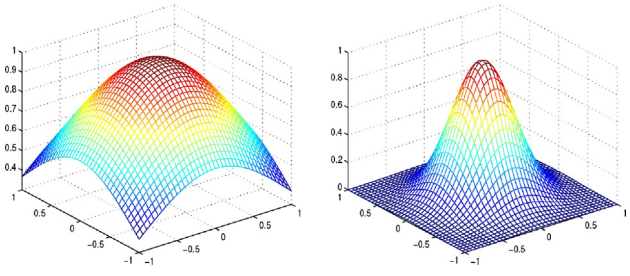

Fig. 2. Gaussian with $r = 1$ (left) and $r = 1/3$ (right) centered at the origin in $\mathbb{R}^2$.

1.1.1. RBF training

There are different learning strategies that we can follow in the design of an RBF network, depending on how the centers of the radial basis functions of the network are specified (Haykin, 1999).

The simplest approach is to assume fixed radial basis functions defining the activation functions of the hidden neurons. The locations of the centers may be chosen randomly from the training data set. In this approach, the only parameter that would need to be learned are the linear weights in the output layer (**W**). The main problem with the method of fixed centers is the fact that it may require a large training set for a satisfactory level of performance (Haykin, 1999).

The centers of the radial basis functions and all others parameter can be defined by supervised methods. Another approach suggests that the training can be separated in two different stages (hybrid learning). The first stage is the self-organized learning, where the purpose is to estimate appropriate locations of the centers (μ) and the spread (σ) of the radial basis functions in the hidden layer. At the second stage, supervised methods completes the design of the network by estimating the linear weights of the output layer (Haykin, 1999).

At the self-organized learning, is utilized a clustering algorithm that partitions the training set into subgroups which should be as homogeneous as possible. A well-know algorithm is the k-means, but any clustering algorithm can be applied at this phase. Algorithm 1 explains how to train a RBF network.

Algorithm 1. RBF Network training

Algorithm 1: RBF Network training

Input Parameters: *in*: input data samples; *out*: output data samples;
Output Parameters: *mse*: mean square error of training; *rbf*: trained network.
begin
  Select centers of radial basis functions using any clustering algorithm;
  for *i=0 to size(in)* **do**
    for *c=0 to size(centers)* **do**
      calculate the output of each input data in RBF function with center in $center_c$;
    end
    Use pseudo-inverse matrix to calculate weights;
    Calculate output of the network with weights adjusted previous;
    Calculate MSE (expected outputs - calculated outputs);
  end
end

There are several clustering algorithms and they can be used to train RBF networks. A recent work (Cruz et al., 2016) demonstrated the feasibility of applying a clustering algorithm inspired by bee behavior to automatically generate the optimal RBF network architecture for data classification. This algorithm has received some changes in order to become more efficient (Silva et al., 2016).

This paper presents RBF networks trained by the bee inspired algorithm as a method to obtain volume estimates of individual trees. The

Table 1
Samples in each forest site.

Site class	Amount of samples
Site I	585
Site II	619
Site III	615
Total	1819

results are compared to the volume equations, RBF networks trained by other methods, and MLP networks.

2. Materials and methods

2.1. Dataset description

The dataset is composed of rigorous scaling data of 1819 trees of a planted forest in the municipality of Aracruz in the State of Espirito Santo - Brazil. 48 circular plots of 360 m^2 were randomly sampled, for a clonal genetic material (hybrid of *Eucalyptus grandis* and *Eucalyptus urophylla*) that is 6.5 years of age, in three (I, II and III) forest sites (16 plots in each). Table 1 details the distribution of amount of samples in each site class. In each of the 48 plots, the DBH (diameter at 1.3 meters from the soil) was measured with a caliper. The total height of the trees (h) of the three central rows were measured with hypsometer (Cabacinha, 2003). Table 2 shows some descriptive statistics of the dataset.

After the forest inventory, all 1819 trees were cut down and rigorously scaled by the Smalian method (Eq. (6)).

$$v_i = l \cdot \frac{(g_1 + g_2)}{2} \tag{6}$$

where v_i is the volume of the section i, l is the length of the section in meters, g_1 and g_2 are the sectional areas of the lower and upper end in square meters. According to Cabacinha (2003), the diameters were measured at 0.1 m, 0.3 m, 0.5 m, 0.7 m, 0.9 m, 1.1 m, 1.3 m, 2 m, 3 m and successively until the total height. The height of each tree was measured separately.

2.2. Evaluated models

In previous works, MLP networks for volume estimation are often used. Therefore, it was decided to consider some architectures of that type of network in this paper as reference. Three architectures were chosen according to Figs. 3–5. These three architectures have each neuron of the intermediate layers implementing the logistic function and those of the output layer implementing the linear function. In all three cases, the training algorithm was the Levenberg-Marquardt with learning rate 0.3 and number of epochs equal to 500. The choice of the training algorithm is justified by the fact of the Levenberg-Marquardt is often the fastest backpropagation algorithm in the toolbox, and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms (Matlab documentation, 2018). All MLP networks were designed and trained in Neural Network Toolbox of Matlab®.

It was also necessary to consider some classical equations as reference. The models of Schumacher-Hall (Eq. (7)) and Spurr (Eq. (8)) were well evaluated in other studies and due to their characteristics like non-biased estimation (Campos and Leite, 2002; Paula Neto, 1977; Campos et al., 1985), these two models were chosen for this work.

Table 2
Descriptive statistics of the data set.

Parameter	Site	Mean	Std.	Min.	Max.
Diameter	III	14.93	2.30	7.15	24.55
	II	15.00	2.29	6.35	21.83
	I	15.20	2.62	4.85	21.75
Height	III	25.65	2.43	14.20	30.60
	II	26.38	2.67	13.00	31.90
	I	27.31	3.47	8.60	33.20
Volume	III	0.22660	0.08248	0.02776	0.64061
	II	0.23992	0.08799	0.02132	0.60083
	I	0.26523	0.10656	0.00833	0.61392

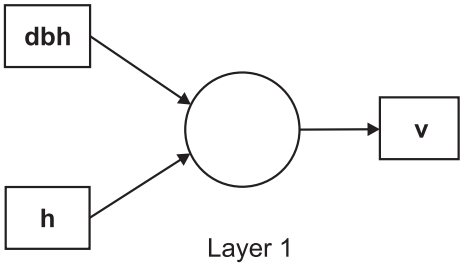

Fig. 3. MLP-1-1.

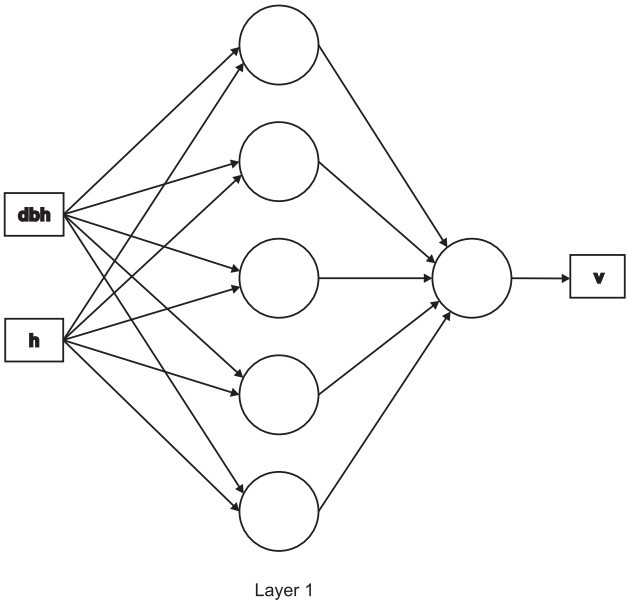

Fig. 4. MLP-1-5.

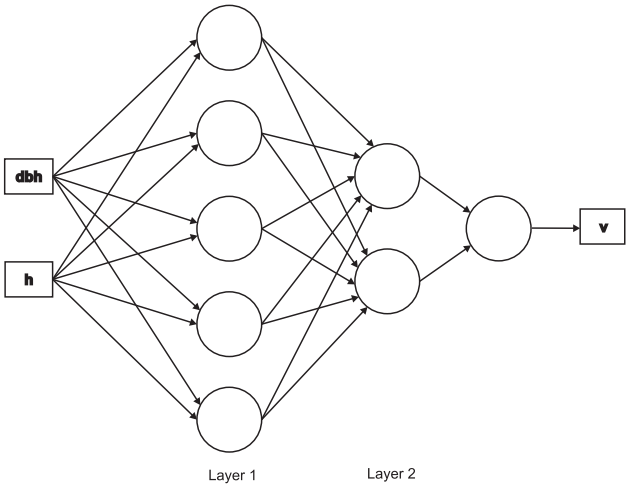

Fig. 5. MLP-2-5-2.

Table 3
Evaluated RBF networks.

Neural Network	Training algorithm	Activation function	Spread heuristic
RBF-A-GU-MAX	random	Univariate Gaussian	maximum
RBF-A-GU-CONST	random	Univariate Gaussian	constant
RBF-A-GU-MED	random	Univariate Gaussian	average
RBF-A-GM	random	Multivariate Gaussian	std. dev.
RBF-A-M-MAX	random	Multiquadric	maximum
RBF-A-M-CONST	random	Multiquadric	constant
RBF-A-M-MED	random	Multiquadric	average
RBF-KM-GU-MAX	k-means	Univariate Gaussian	maximum
RBF-KM-GU-CONST	k-means	Univariate Gaussian	constant
RBF-KM-GU-MED	k-means	Univariate Gaussian	average
RBF-KM-GM	k-means	Multivariate Gaussian	std. dev.
RBF-KM-M-MAX	k-means	Multiquadric	maximum
RBF-KM-M-CONST	k-means	Multiquadric	constant
RBF-KM-M-MED	k-means	Multiquadric	average
RBF-BEE-GU-MAX	cOptBees	Univariate Gaussian	maximum
RBF-BEE-GU-CONST	cOptBees	Univariate Gaussian	constant
RBF-BEE-GU-MED	cOptBees	Univariate Gaussian	average
RBF-BEE-GM	cOptBees	Multivariate Gaussian	std. dev.
RBF-BEE-M-MAX	cOptBees	Multiquadric	maximum
RBF-BEE-M-CONST	cOptBees	Multiquadric	constant
RBF-BEE-M-MED	cOptBees	Multiquadric	average

$$v = \beta_0 \cdot dbh^{\beta_1} \cdot ht^{\beta_2} \tag{7}$$

$$v = \beta_0 + \beta_1 \cdot dbh^2 \cdot ht \tag{8}$$

The main objective of this paper is to evaluate the RBF performance, so different characteristics and training algorithms were chosen for this network. Three training algorithms were chosen for the unsupervised phase, three activation functions for the intermediate neuron layer and three heuristics for the definition of the spread of these functions. Table 3 shows the names adopted for each RBF network and their respective properties.

In networks with random architecture, the number of neurons and their respective centers were randomly defined by a uniform distribution, in which the number of neurons can vary between 2 and half of the number of samples used in the training. For the architecture defined by the k-means algorithm, the number of clusters was empirically set as 10.

The version of cOptBees (Algorithm 2) used was the one proposed by Silva et al. (2016), parameterized according to Table 4.

Algorithm 2. cOptBees

Table 4
cOptBees Parameters.

Parameter	Value
Max clusters	$N/2$
Min bees	50
Max bees	100
Percentage of recruiters	90%
Recruitment rate	0.7
Recruited for each recruiter	5
Inhibition radius	0.002
Stop criterion	5 cycles without fitness evolution

Algorithm 2: cOptBees

Input Parameters: n_{min}: minimum number of active bees; n_{max}: maximum number of active bees; ρ: inhibition radius; α: recruitment rate; ε: foraging effort; p_{min}: minimum probability of a bee being a recruiter; p_{rec}: percentage of non-recruiters that will be actually recruited; $rMax$: maximum number of clusters.

Output Parameters: Active bees and their respective values of objective function.

begin
  Ramdomly generate a swarm of n bees;
  while *stopping criterion is not attained* **do**
    Evaluate the quality of the sites being explored by the active bees;
    Apply local search;
    Determine the recruiter bees;
    Update the number of active bees;
    Determine the recruited and scout bees;
    Perform the recruitment process;
    Perform the exploration process;
  end
end

In order to define the spread of the radial base functions, three heuristics were used: constant, average, and maximum. The heuristic called constant defines the spread of all functions equal to 1. The heuristic that has been called maximum consists in using the Eq. (9) in the calculation of the spreads

$$\sigma_1 = \sigma_2 = ... = \sigma_n = \frac{d_{max}}{\sqrt{2}} \tag{9}$$

where n is the amount of centers and d_{max} is the biggest Euclidean distance between all centers (Bao et al., 2011). The heuristic that received the name of average define the spread of each function as the average of Euclidean distances between the current center to all other centers (Eq. (10))

$$\sigma_k = \frac{1}{n-1} \sum_{i=1, i \neq k}^{n} ||c_k - c_i|| \tag{10}$$

where σ_k is the spread of the cluster k, whose center is c_k and n is the total number of centers (Chen et al., 2008).

The supervised training phase, where the synaptic weights are adjusted, for all evaluated RBF's, was done by the calculus of the pseudo-inverse matrix. The equations and RBF models were coded in Matlab®.

2.3. Performance metrics

In order to know the accuracy of the estimations provided by each technique, a traditional and a "inverse" k-fold cross-validation were done, both with number of folds equal to 10 ($k = 10$). The performance

metrics utilized here was the Mean Absolute Percentage Error (MAPE) (Eq. (11)) because it is easy to interpret and is independent of the variable scale. As this measure represent the estimation error, is desirable that his value are as close to zero as possible.

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{|Y_i - \hat{Y}_i|}{|Y_i|} \quad (11)$$

where n is the number of samples, Y is the observed value and $\hat{Y}$ is the estimated value.

Although these performance metrics are an important indicator of the quality of the solution provided by a given model, other properties must also be taken into account. Among these properties, one can list the scalability and the need for knowledge about the data.

In addition to the performance metrics, the models were evaluated using hypothesis testing to verify if there was any indication that the differences between the results were significant.

3. Results and discussion

The dataset was evaluated according to the site class to which they belong. Thus, 3 k-fold cross-validations were performed, one for each site.

Initially, in an attempt to find models that could not accurately represent the volumes of the trees, the estimated volumes by each model in the cross-validation were computed and compared to the observed volumes. Before comparing the models, a normality test was performed to aid in the decision regarding the use of parametric or non-parametric test.

To verify the normality of the estimated volumes, the normality test of Lilliefors was used with a significance level of 5%. The Lilliefors test, according to Cirillo and Ferreira (2003), is quite flexible in that the null hypothesis H_0 specifies that the population belongs to the normal distribution family, without however having to specify the mean or variance of the distribution. This test had its null hypothesis rejected in at least one case, which indicates that not all estimated volumes follow a distribution of the normal distribution family.

As it was not possible to expect normality in all data, it was chosen to perform non-parametric tests to compare the results. The Wilcoxon test with a significance level of 5% was used in the peer-to-peer comparison between the observed volume and the estimated volume by each of the 26 methods. In this evaluation, the test didn't have its hull hypothesis H_0 reject in any case, which suggests that, for the significance level of 5%, there are no significant differences between the observed volumes and the estimated volumes by each of the models. Therefore, this preliminary assessment did not allow any model to be disregarded.

The previous analysis does not suggest the elimination of any model, so all models were submitted to residual analysis. Table 5 presents the MAPE obtained by each model in the volume estimate of individual trees of the site class.

In this table, it is observed that some models presented very high MAPE, which is justified by the presence of outliers in the estimation promoted by them.

To evaluate the quality of the estimate provided by each model, they were grouped into 5 classes: equations, MLP networks, RBF random networks, RBF k-means and RBF cOptBees. This grouping was designed to simplify the comparison, because instead of comparing 26 models, only 5 are compared, with one model representing each class of models. To make this reduction possible, it was necessary to adopt a criterion of choice of the best model to represent each class. As the mean of the error was distorted due to the presence of outliers, the median was adopted as the criterion of choice, since its measure is more robust to outliers. Thus, the best model of each class is the one that has the modulus of the median closest to zero. The results of the best models of each class were compared using hypothesis tests.

Table 5
Mean Absolute Percentage Error (MAPE) of each model.

Id.	Model	site I	site II	site III
1	Schumacher-Hall	4.55	**3. 76**	4.26
2	Spurr	5.01	4.09	4.49
3	MLP-1-1	4.61	3.92	4.48
4	MLP-1-5	8.44	3.84	**4. 13**
5	MLP-2-5-2	**4. 33**	3.77	4.80
6	RBF-A-GU-MAX	6.87	4.42	31.55
7	RBF-A-GU-CONST	4.62	4.46	8.37
8	RBF-A-GU-MED	10.27	4.48	93.62
9	RBF-A-GM	4.62	4.20	11.05
10	RBF-A-M-MAX	5.28	4.70	13.73
11	RBF-A-M-CONST	4.69	4.64	8.98
12	RBF-A-M-MED	5.81	5.03	6.48
13	RBF-KM-GU-MAX	5.54	4.98	6.32
14	RBF-KM-GU-CONST	4.63	4.47	8.36
15	RBF-KM-GU-MED	5.38	4.83	7.71
16	RBF-KM-GM	4.71	4.04	4.75
17	RBF-KM-M-MAX	5.41	4.94	14.14
18	RBF-KM-M-CONST	4.71	4.71	8.95
19	RBF-KM-M-MED	5.92	4.14	16.43
20	RBF-BEE-GU-MAX	8.83	4.49	4.85
21	RBF-BEE-GU-CONST	4.63	4.46	8.37
22	RBF-BEE-GU-MED	10.21	4.57	23.51
23	RBF-BEE-GM	5.21	4.29	9.66
24	RBF-BEE-M-MAX	6.69	4.82	12.89
25	RBF-BEE-M-CONST	4.70	4.66	8.98
26	RBF-BEE-M-MED	7.64	4.85	8.04

3.1. Results for k-fold cross-validation

3.1.1. Results for site I

For site I data, the Wilcoxon test indicated that the results of the two equations are equivalent. Because of the smaller modulus of the median, the Spurr equation was chosen as representative of the classical models. Also according to this test, the three MLP's are equivalent, but due to the median error, the MLP-1-1 was chosen as representative of the MLP's. For RBF's with random choice of centers, the lowest median was RBF-AM-CONST (the statistical test showed that this model differs only in its class from RBF-A-GU-CONST and RBF-AM-MED). For RBF networks using the k-means algorithm, RBF-KM-GU-MAX, which differs significantly only to the RBF-KM-GM and the RBF-KM- M-MED. For RBF trained with cOptBees, the one with the best median error was the RBF-BEE-M-CONST, being statistically different only from the RBF-BEE-M-MED. Table 6 shows the median values of the residuals of the models selected as representative of each class.

After choosing the representatives, their residues were submitted to the statistical test to evaluate the relevance of the difference between them. The RBF-KM-GU-MAX model was the model with the median closest to zero. The results were compared to the others using the Wilcoxon test with a significance level of 5%, which indicated that the three types of RBF are statistically equivalent to each other and differ from equations and MLP's.

To provide more information about the equivalence between the

Table 6
Median of errors in volume estimation in site I.

Id.	Model	Median
2	Spurr	9.66×10^{-5}
3	MLP-1-1	-1.65×10^{-4}
11	RBF-A-M-CONST	-8.03×10^{-5}
13	RBF-KM-GU-MAX	-3.09×10^{-5}
25	RBF-BEE-M-CONST	-6.31×10^{-5}

Table 7
Cases where the null hypothesis of the Wilcoxon test was rejected in site I.

Id.	2	3	11	13	25
2				•	
3			•	•	•
11		•			
13	•	•			
25		•			

models, Table 7 was elaborated, in which the five models are arranged in rows and columns and, if there is rejection of the null hypothesis of the test between a certain pair of models, the symbol (•) appears at the corresponding row × column intersection.

3.1.2. Results for site II

The same process of choice was carried out in the data of site II. In this case, the models chosen based on the median criteria were: Schumacher-Hall, MLP-1-1, RBF-A-GU-CONST, RBFKM-GU-MAX, and RBF-BEE-M-MAX. Statistical tests indicated that the chosen models are statistically equivalent to the other members of their classes, except for random RBF's, in which RBF-A-GU-CONST differs from RBF-A-GU-MAX and RBF-AM-CONST. The median error values of each model can be seen in Table 8.

Among the models chosen according to the median criterion, the model that presented a median closer to zero was the RBF-KM-GU-MAX and, according to Wilcoxon's test, it statistically differs only from RBF-A-GU-CONST. The Table 9 indicates the other cases in which rejection of the null hypothesis of the Wilcoxon test occurred for the data of site II.

3.1.3. Results for site III

For the data from site III, according the criterion of the median of the error, the models Spurr, MLP-1-1, RBF-A-GU-CONST, RBF-KM-GU-CONST, and RBF-BEE-GU-CONST were chosen as representatives of the each class of models. In the equations class, both equations are statistically equivalent and the same fact occurs in relation to MLP's. In the random RBF's class, the selected RBF differs from the RBF-A-M-CONST; in the class of k-means RBF's, the chosen RBF differs only from the RBF-KM-M-CONST; in the class of RBF's trained by cOptBees there is statistic difference only between the chosen model and the RBF-BEE-M-CONST. In the Table 10 the median of the error of each chosen model can be observed.

At the inter-class comparative, the RBF-BEE-GU-CONST presents median closest to zero, however their results, according the Wilcoxon test, are statistically equivalent to all others.

3.2. Results of inverse k-fold cross-validation

The previous experiment used the standard cross-validation to evaluate the performance of each model. However, in this type of validation, most of the data (90%) is used for training and a small portion (10%) is used for validation. For the forest inventory, its is ideal to use models that can be adjusted with a small number of samples. Therefore, the models were submitted to a reverse cross-validation (Treiber et al., 2012), that is, instead of training with 90% of the data and validate with 10%, the models were trained with 10% and validated with 90%.

Table 8
Median of errors in volume estimation in site II.

Id.	Model	Median
1	Schumacher	1.50×10^{-4}
3	MLP-1-1	8.10×10^{-5}
7	RBF-A-GU-CONST	1.80×10^{-4}
13	RBF-KM-GU-MAX	7.47×10^{-6}
24	RBF-BEE-M-MAX	5.75×10^{-5}

Table 9
Cases where the null hypothesis of the Wilcoxon test was rejected in site II.

Id.	1	3	7	13	24
1		•			•
3	•				•
7				•	
13			•		
24	•	•			

Table 10
Median of errors in volume estimation in site III.

Id.	Model	Median
2	Spurr	4.47×10^{-5}
3	MLP-1-1	5.93×10^{-4}
7	RBF-A-GU-CONST	1.89×10^{-5}
14	RBF-KM-GU-CONST	3.46×10^{-5}
21	RBF-BEE-GU-CONST	5.71×10^{-6}

After performing all validation steps for the three sites, the estimated volumes were compared to the observed volumes using the Wilcoxon test at a significance level of 5%. Table 11 shows the models whose estimated volumes differ significantly from the observed volumes at each of the three sites.

3.2.1. Results for the data from the site I

After eliminating the models that could not correctly estimate the volumes, the others were evaluated according to the same methodology used in the previous section. The models with lower modulus of the median of the error can be observed in Table 12.

The model with the best median was the equation of Spurr. Their results were compared to the others using the Wilcoxon test. The test results indicated that the equations, RBF and MLP's trained cOptBees are equivalent to each other, but were statistically different from the random RBF and the k-means RBF. Table 13 details the test results, highlighting the line intersection × column in which the null hypothesis of the test was rejected.

3.2.2. Results for the data from the site II

According to Table 11, two MLP's and some randomly generated RBF's and through cOptBees were not able to estimate the volumes correctly. Thus, these models were disregarded and the others were analyzed according to the median criterion. Table 14 shows the median error values of the selected models.

The model with the best median was the RBF-KM-GM, according to the results of the Wilcoxon test. This network is equivalent to the RBF-BEE-GU-MED and different from the other models. In Table 15, it is possible to observe all the cases in which there was rejection of the null hypothesis of the test.

Table 11
Models that were unable to estimate volumes correctly.

site I	site II	site III
RBF-A-GU-MAX	MLP-1-5	
RBF-A-GU-CONST	MLP-2-5-2	MLP-1-1
RBF-A-M-MAX	RBF-A-GU-MED	RBF-A-GU-MED
RBF-A-M-CONST	RBF-A-M-MAX	RBF-KM-GU-MED
RBF-A-M-MED	RBF-BEE-GU-MAX	RBF-KM-GM
RBF-BEE-GU-MAX	RBF-BEE-GM	RBF-BEE-GU-MAX
RBF-BEE-GU-MED	RBF-BEE-M-MAX	
RBF-BEE-M-MED	RBF-BEE-M-MED	

Table 12
Median of errors in volume estimation in site I.

Id.	Model	Median
2	Spurr	2.01×10^{-5}
3	MLP-1-1	-2.56×10^{-4}
8	RBF-A-GU-MED	5.47×10^{-4}
19	RBF-KM-M-MED	-1.66×10^{-4}
23	RBF-BEE-GM	-6.00×10^{-4}

Table 13
Cases where the null hypothesis of the Wilcoxon test was rejected in site I.

Id.	2	3	8	19	23
2			•	•	
3			•	•	
8	•	•			•
19	•	•			•
23			•	•	

Table 14
Median of errors in volume estimation in site II.

Id.	Model	Median
1	Schumacher	1.41×10^{-4}
3	MLP-1-1	8.04×10^{-4}
9	RBF-A-GM	1.06×10^{-4}
16	RBF-KM-GM	2.86×10^{-5}
22	RBF-BEE-GU-MED	-1.11×10^{-4}

Table 15
Cases where the null hypothesis of the Wilcoxon test was rejected in site II.

Id.	1	3	9	16	22
1			•	•	
3			•	•	
9	•	•		•	
16	•	•	•		
22					

3.2.3. Results for the data from the site III

As it was done with the data from sites I and II, disregarding the models that, according to an initial analysis, could not correctly estimate the volumes of the trees, for the data from site III, the remaining models were chosen according to the median of the residue. These models and their respective medians can be observed in Table 16.

After selecting the models according to the criterion of the median, they were compared to each other by means of the Wilcoxon test. Among the five models selected, the one that presented the lowest median was the RBF-A-GU-CONST. The test results indicated that all models differed significantly from each other.

Table 16
Median of errors in volume estimation in site III.

Id.	Model	Median
1	Schumacher	6.11×10^{-5}
5	MLP-2-5-2	-4.10×10^{-4}
7	RBF-A-GU-CONST	3.74×10^{-5}
18	RBF-KM-M-CONST	2.08×10^{-4}
23	RBF-BEE-GM	1.36×10^{-4}

4. Conclusion

This study aimed to evaluate the RBF neural networks generated by cOptBees algorithm as a way to estimate volume of individual trees. In order to evaluate the performance of RBF networks, their results were compared to those of the classical equations and MLP networks applied to the same database of eucalyptus volumes.

There are several equations which can be used to estimate the volume of trees. Among these, two were chosen as reference for being well evaluated in previous studies. MLP networks have already been object of study in several works of forestry. For that reason, this study also had to consider some configurations of MLP as a comparison parameter. There is no standard method for defining MLP network parameters based on the problem data. So, three different MLP configurations have been chosen to be confronted to the RBF's. For the RBF networks, three training algorithms with different characteristics, three activation functions and three heuristics were chosen to define the spread of these functions to observe the effect that these factors would cause on the results.

26 models were chosen to be evaluated in data from three forest sites. These 26 models comprise 2 equations, 3 MLP's and 21 RBF's. All models were submitted to two types of 10-fold cross-validation: standard and inverse. The data set is composed of 1819 trees distributed in three site class and the forests sites (I, II, and III) which have, respectively, 585, 619, and 615 samples. At the 10-fold cross-validation, 9 folds are used to training and 1 fold to validation, which is, approximately, 500 samples to training and 50 to validation. At the inverse 10-fold cross-validation, occurs the inverse situation. The results obtained make it possible to conclude that when there was a high availability of data, around 500 samples, to fit the models, all of them were able to estimate the volumes in a similar way. This is because when the models were evaluated by the standard cross-validation method, about 500 samples were used for training and 50 for validation at each forest site. It was observed that all variations of MLP's and RBF's were also able to estimate the volumes, which indicates that when there is a lot of data for training, any configuration of these models has equal capacity to estimate volumes. None of the RBF training algorithms has been able to stand out from the others in terms of the quality of results.

When the amount of samples available for training was reduced to about 50 samples, not all models were able to correctly estimate tree volumes. The number 50 comes from reverse cross-validation that uses only 10% of the data for training and the sites have around 500 sample trees. Only the classical models were not influenced by this large variation in the amount of training data as they were able to estimate the volume without significant differences in all evaluated situations.

The MLP's and the RBF's were strongly affected by the lack of data for training, because in some cases in the reverse cross-validation, the statistical test indicated significant differences between the volumes that were estimated and the observed volumes. From the training algorithms evaluated, only cOptBees allows the RBF to achieve similar results to the classic models and the MLP's, which shows that the RBF that is generated randomly or by k-means can produce unsatisfactory results.

Therefore, it is not possible to consider that the use of neural networks for the estimation of volumes represented an improvement in the accuracy for the database studied in this study, because in all performed experiments no differences were observed in favor of the neural networks, MLP or RBF. Not even the use of a more elaborate algorithm for the training of the RBF's was enough to make them better than the classic models.

Therefore, what this technique brings as a benefit is its more general structure, able to adapt to the different data sets and produce acceptable estimates, thus avoiding the need for the forest engineer to adjust several mathematical models to later choose the best one. RBF networks, in particular, have the advantage of having architecture defined automatically by clustering algorithms, a fact that makes them easier to

work with when compared to MLP's.

Although the results of this study did not indicate that neural networks provide better estimates than classical methods, it is important to remember that only one database was considered as tree samples. As future study, would be possible to propose the comparison of the neural networks to the classical methods in another database in which there is a higher shape variability or even using another species of tree to verify the generalization capacity of each one of the methods in a more challenging situation. We are developing a user-friendly interface software to the presented methodology. So, other peoples may apply our algorithm to solve practical problems without needing to have deep knowledge about RBF or clustering algorithms.

References

Bao, P., Yuan, M., Song, H., Guo, W., Xue, J., 2011. Aircraft wing structural damage localization research based on rbf neural network. In: 2011 IEEE 5th International Conference on Cybernetics and Intelligent Systems (CIS), Quingdao - China, pp. 57–62.

Binoti, D.H.B., Binoti, M.L.M.S., Leite, H.G., 2014. Configuracao de redes neurais para estimacao de volume de arvores. Ciencia da Madeira.

Blanco, A.M., Sotto, A., Castellanos, A., 2013. Prediction of amount of wood using neural networks. J. Math. Modell. Algorithms Oper. Res. 11, 295–307.

Cabacinha, C.D., 2003. Um metodo para a realizacao do inventario florestal suprimindo a cubagem rigorosa. Mestrado em engenharia florestal, Universidade Federal de Lavras, Lavras - Minas Gerais.

Campos, J.C.C., Leite, H.G., 2002. Mensuracao Florestal: perguntas e respostas, UFV, Vicosa.

Campos, J.C.C., Trevisol Jr., T.L., Paula Neto, F., 1985. Ainda sobre a selecao de equacoes de volume. Revista Arvore 9, 115–126.

Chen, J.Y., Qin, Z., Jia, J., 2008. A pso-based subtractive clustering technique for designing RBF neural networks. In: Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence). IEEE Congress on, Hong Kong, pp. 2047–2052.

Cirillo, M.A., Ferreira, D.F., 2003. Extensao do teste de normalidade univariado baseado no coeficente de correlacao quantil-quantil para o caso multivariado. Revista de Matematica e Estatistica 21, 67–84.

Cruz, D.P., Maia, R., Silva, L.A., Castro, L.N., 2016. Beerbf: a bee-inspired data clustering approach to design rbf neural network classifiers. Neurocomputing 427–437.

Gorgens, E.B., Leite, H.G., Santos, H.N., Gleriani, J.M., 2009. Estimacao do volume de arvores utilizando redes neurais artificiais. Revista Arvore.

Haykin, S., 1999. Neural Networks - A Comprehensive Foundation, second ed. Prentice Hall, NJ - USA.

IBA, 2017. Report IBA - brazilian tree industry, http://iba.org/images/shared/Biblioteca/IBA_RelatorioAnual2017.pdf. (online, accessed 2017-12-03).

Lacerda, T.H., Cabacinha, C., Junior, C.A., Maia, R., Lacerda, K.W., 2018. Artificial neural networks for estimating volume of trees in cerrado. CERNE 23, 483–491.

Matlab documentation, 2018. <https://www.mathworks.com/help/nnet/ref/trainlm.html>. (online, Accessed 2018-04-29).

Ozcelik, R., Diamantopoulou, M.J., Brooks, J.R., Wiant Jr., H.V., 2010. Estimating tree bole volume using artificial neural network models for four species in turkey. J. Environ. Manage. 91, 742–753.

Ozcelik, R., Diamantopoulou, M.J., Brooks, J.R., 2014. The use of tree crown variables in over-bark diameter and volume prediction models. iForest - Biogeosciences Forestry 132–139.

Paula Neto, F., 1977. Tabelas volumetricas com e sem casca para *Eucalyptus grandis*. Revista Arvore 1, 31–54.

Pazouki, M., Wu, Z., Yang, Z., Mller, D.P.F., 2015. An efficient learning method for RBF neural networks. In: 2015 International Joint Conference on Neural Networks (IJCNN), pp. 1–6.

Sanquetta, C.R., Piva, L.R., Wojciechowski, J., Corte, A.P., Schikowski, A.B., 2018. Volume estimation of cryptomeria japonica logs in southern brazil using artificial intelligence models. Southern Forests: J. Forest Sci. 80, 29–36.

Silva Jr., E.M., Maia, R.D., Cruz, D.P.F., Castro, L.N., 2016. Improvements for coptbees clustering algorithm, Vitoria - ES - Brazil, pp. 803–808.

Treiber, M., Kesting, A., Thiemann, C., 2012. Traffic Flow Dynamics: Data, Models and Simulation. Springer, Berlin Heidelberg.

Wang, S.-H., Zhang, Y., Li, Y.-J., Jia, W.-J., Liu, F.-Y., Yang, M.-M., Zhang, Y.-D., 2018. Single slice based detection for alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. Multimedia Tools Appl. 77, 10393–10417.

Zhang, Y., Sun, Y., Phillips, P., Liu, G., Zhou, X., Wang, S., 2016. A multilayer perceptron based smart pathological brain detection system by fractional fourier entropy. J. Med. Syst. 40, 173.