Use of artificial neural networks to assess yield projection and average production of eucalyptus stands

Aline Edwiges Mazon de Alcântra¹, Ana Carolina de Albuquerque Santos¹*, Mayra Luiza Marques da Silva², Daniel Henrique Breda Binoti³, Carlos Pedro Boechat Soares¹, José Marinaldo Gleriani¹ and Helio Garcia Leite¹

¹Department of Forest, Universidade Federal de Viçosa, CEP 36570-000, Viçosa – MG, Brazil.
²Department of Forestry, Alto Universitário, Universidade Federal do Espírito Santo, Guararema, CEP 29500-000, Alegre/Espírito Santo, Brazil.
³DAP Florestal, R. Papa João XXIII, 9 - CEP 36570-000, Viçosa – Minas Gerais, Brazil.

Received 19 December, 2017; Accepted 11 May, 2018

Eucalyptus stands growth depends on genotype, age, quality of the local soil and silvicultural treatment. Environmental factors, mainly the water availability to plants throughout the years, temperature and solar radiation are relevant to production capacity. The models used in Brazil to stimulate the future production of forestry stands are those that estimate growth and/or production according to age, basal area and local index. One of the possible approaches to do so is the use of procedural models (ecophysiological) such as the 3PG and the artificial neural network. The current study has the aim to construct, validate and apply an artificial neural model to predict the production and growth of eucalyptus stands in Minas Gerais, Brazil. The herein used data resulted from continuous forestall inventory plots conducted in eucalyptus stands in the North, Center and South of the state. The edaphic and climatic information added to the IFC data were used to train neural nets on predicting growth and production in the state. A neural network, lacking inventory variables, was also trained to extrapolate the mean productivity in the entire state of Minas Gerais due to the physiographic, edaphic and climatic conditions. The neural network efficiency was attested by the great accuracy of productivity forecasts. The generated productivity maps are indicated for studies on the expansion of eucalyptus cultivation in the state. The applied methodology is simple and efficiently inapplicable to different forestry cultures in other states or countries.

Key words: Eucalyptus, water availability, forestry stands, neural network, productivity.

INTRODUCTION

The Forestry Sector is responsible for approximately 3.5% of Brazilian Gross Domestic Product (2007 GDP) and it accounts for 7% of total exportation, thus generating 7 million jobs (Brazilian Forestry Service, 2015). From the 8 million hectares of planted forest, about 5.6 million are Eucalyptus spp. stands basically located in the states of Minas Gerais (25.2%), São Paulo (17.6%) and Mato Grosso do Sul (14.5%) (IBA, 2015). The eucalyptus stands are composed of many clones (Gonçalves et al., 2008) and spatial arrangements conducted in high stem and coppice systems. The timber is mainly used in cellulose and charcoal production
(Campos and Leite, 2013). The cutting time in Brazil takes place when the trees are around seven years old and the mean production (38 m³ ha⁻¹) varies a lot (IBA, 2014). Production may potentially reach 90 m³ ha⁻¹ when the tree is at the age of 6 (Borges, 2012). In 2014, cellulose and charcoal production in Brazil was about 80,873,295 million m³ and 6,219,325 tons (IBGE, 2014). The timber is also used in the construction industry, for electric power generation, and in plates (IBA, 2015). The harvest often occurs approximately five to eight years after the plantation or after the sprouting. Yet, the harvest period depends on the regulation model applied by the company, that is, on the hierarchical planning.

Minas Gerais state has the largest eucalyptus plantation area in Brazil. These plantations are mostly located in regions called Mares de Morros, Tabuleiros Costeiros (Gonçalves et al., 2008), and Brazilian Savanna (Cerrado), wherein the soil fertility is relatively low and the water availability is irregular throughout the years (Lelles et al., 2001). The mean productivity of seven-year-old stands (cutting point) varies according to physiographic, edaphic and climatic conditions. It also changes according to genotype, spacing and spatial arrangements, the productive capacity of the area, cultural practices (Stape et al., 2006, 2010; Oda-Souza et al., 2008), silvicultural management (Gonçalves et al., 2008) and water availability regularity (Ryan et al., 2010 and Stape et al., 2010).

Most of the Brazilian eucalyptus stands belong to more than 6,000 forestry companies in the country and to forestry smallholders supported by promotion programs (IBA, 2015). The study was done in partnership with researchers from 66 current forestry engineering courses placed in the country in 2015 (SNIF, 2016).

Research institutions, universities and engineers from forestry companies have been conducting studies on the growth and production of eucalyptus stands for many years in Brazil, developing large databases of permanent plot and tress cubage samples. These data are mostly used by companies assisted by researchers from public institutions and in partnership with universities to develop growth and yield models (GYM). The most used GYMs in Brazil are the total stands types (Campos and Leite, 2013). Trevizol (1985) published one of the first consistent variable density models for Amazonian eucalyptus stands. Thenceforth, many studies were individually conducted by different companies, most of these studies focused on defining stratification due to cutting process, spatial arrangement, genotype and region or project (Trevizol, 1985; Soares, 2000; Cruz et al, 2008; Nogueira, 2005; Oliveira et al., 2009; Silva, 2010; Borges, 2009; Salles, 2012; Binoti, 2012). Campos and Leite (2013) published a summary of the main models and functional relationships used in Brazilian eucalyptus stands. According to the authors, the most commonly applied models are the sigmoid and Clutter’s (1963). Larger companies have been following the modeling approaches that present the maximum forest stratification and, consequently, the sigmoid prediction models such as the logistic Gompertz (Winsor, 1932) and Richards (Richards, 1959) ones, despite the other variables resulting from the Von Bertalanffy (1938) model. The functional relationships used are the types: \( V = f(I), \) \( V = f(I,S) \) or \( V = f(I,S,B) \), wherein \( V \) is the production per hectare, \( B \) is the basal area per hectare, \( S \) is the local index and \( I \) is the age of the stands.

Besides the large range of physiographic, edaphic and climatic feature effects, one of the challenges faced by the production and yield modeling of Brazilian eucalyptus is the large diversity of genotypes (mainly clones) that interact in time and space with different spacing, spatial arrangements and handling types. The constant genotypes often change and the silvicultural practices hamper the modeling. It is common to find 30 to 60% of permanent sample parcels with just one or two calibrations (Oliveira et al., 2009) and results from current technological packages are often quite important for the companies (Oliveira et al., 2009; Campos and Leite, 2013).

The Brazilian eucalyptus is also hard to model due to the old databases that are often discarded because of the new “technological packages” that have been employed lately (Oliveira et al., 2009). A statistician would say: “it is impossible to develop a biologically consistent model only by counting on one or two measurements of permanent parcels”. Nevertheless, the problem is incredibly challenging, since there is a large variety of calibrations applied to permanent parcels, there are handling unities or projects composed of none, one, two, three or more calibrations, there is the use of old places, where the plantation has not been initiated WITH the use of various technological packages (Oliveira et al., 2009). It is necessary to attend to the third handling element in the hierarchical plan of eucalyptus forests (Campos and Leite, 2013); having harvest storage expectation, even for areas where the crop that will possibly contain new genotypes and that will be planted in future years; and also counting on expectations on a whole spectrum of planning that encompasses fifteen to thirty years ahead.

Brazilian researchers are very interested in the influence of climatic changes on the development and production of eucalyptus stands countrywide, in remarkable and sometimes uncertain ways (Baesso et
al., 2010). Thereby, predictions are getting harder with time, since the modeling process is often based on past data, due to the assumption that environmental conditions will repeat themselves, but the fact is that the environmental conditions are increasingly uncertain and their inclusion in the models is not trivial (Soares, 2000). New technological packages are implemented in a yearly basis and the cultivated areas are extended to different places that present other physiographic, edaphic and climatic features.

Despite the modeling complexity, it is necessary having accurate production estimations, since the hierarchy planning depends on it. A possibility to improve production estimation accuracy in comparison with the usual production and yield models lies on the computational intelligence methods (CI) such as the artificial neural networks (ANN) (Binoti, 2012). The CI methods used in forestry sciences have been largely employed in the forestry calibration and in pattern classification areas. The Multilayer Perceptron model (MP) is often used for logistical activation (Guan and Gertner, 1991a, b; 1995; Gordon, 1998; Diamantopoulou, 2005; Silva, 2010; Binoti, 2012; Özçelik et al., 2010; Özçelik et al., 2013; Diamantopoulou, 2010a, b; Khoury Junior et al., 2006; Leduc et al., 2001; Silva et al., 2009, Leite et al, 2011; Binoti et al, 2014, b; Gorgens et al., 2009). Most studies employing ANN to predict the Brazilian eucalyptus stands have been using specific data from certain companies or locations (Silva, 2009; Binoti, 2012; Binoti et al., 2012). The ANN allows assessing and/or simulating the climatic, edaphic and silvicultural effects on the productivity of the stands; although it looks superior in many studies when it is compared with current methods, which were described in most of the herein mentioned researches. It also allows estimating the production and productivity in areas that do not have trees or inventory data available (Silva Binoti, 2012).

The production and yield models are employed not only to the forestry handling management of Brazilian eucalyptus, but in certain cases, to ecophysiological models such as the “Physiological Processes Predicting Growth” (3PG) (Sands and Landsberg, 2002; Miehle et al., 2009), which have been already tested and applied in Brazil (Stape, 2002; Almeida et al., 2003, 2004; Stape et al., 2004; Borges, 2009; Stape et al., 2010; Borges, 2012). These models describe the stands growth based on processes linked to physical (soil and climate) and biological features (genetic materials and plant physiology). The aforementioned models are efficient to assess productivity losses caused by root issues and to determine the potential productivity. However, they demand data from directed trial and stands samples. On the other hand, the ANN may be trained through the employment of parcel data from continuous and temporary forestry inventories, as well as through the use of different climatic, edaphic and physiographic calibrations, which are somehow an alternative to the processual models.

It is worth considering the combination of features that express these factors when the MCP is adjusted, since the trees depend on climatic, edaphic and physiographic factors, as well as in genotype, spacing and silvicultural practices to develop. However, the inclusion of categorical variables such as soil, name of the project, genotype, spacing, fertilization level, among other categorical variables, is not trivial and it is often impossible to be done due to lack of representativeness of all the combined variables. It may bring representativeness deficiency to some strata, and it would lead to the need for and to the possibility of using the CI and ecophysiographic models. The possibility of including any continuous or categorical variable and the fact that it is not necessary to observe the statistical preconditions of the regression modeling are some of the advantages of applying the IC model, which does not need a large amount of data from the categorical combinations (Jensen et al., 1999).

The aim of the present study is to develop and validate artificial neural network models in order to identify the production and development of eucalyptus stands in the State of Minas Gerais by taking into account its great potential to produce eucalyptus and its importance to the Brazilian economy. Another aim of the current study is to reduce the investment on these new kinds of plantations in the state by configuring, training, validating, and applying neural networks through the generation of productivity category maps; and to define a methodology to be applied in other states or countries.

MATERIALS AND METHODS

Data

Data from permanent parcels of continuous forestry inventories (CFI) were used in the current study, which was conducted in eucalyptus stands in Minas Gerais – Brazil (Figures 5 and 6). The studied stands are located in North, Center and South of the state; they result in 10,000 parcels that encompass an area of approximately 500 m² aged 12 to 357 months and account for 317,000 registers in the database. All the recorded information and six hierarchical area division levels were standardized—the smallest unit has a 20 hectares edge. The registration form listed city, plantation date, spacing, genotype, predominant soil and rotation. The IFC data were processed according to the parcel level and they comprised all the variables, such as age (months), mean height of predominant trees (Hd), basal area (B) and trees with commercial bark volume (diameter equals to or greater than 4 cm) (Table 1).

Physiographic, edaphic and climatic information from the climatic stations were added to the database, besides the forestry inventory data (Table 2). The annual mean of physiographic, edaphic and climatic information from 2006 to 2013 were processed. The data from the climatic station where connected according to their geographic coordinates and processed in the database in order to extrapolate the information of each edge using the Thiessen polygons methodology (Thiessen, 1911). The climatic information was obtained through Köppen Geiger classification, which takes
into account the seasonality and the annual and monthly mean temperature and precipitation (Köppen and Geiger, 1928).

### Developing and applying neural networks

Two studies were conducted. The aim of the first (Case 1) was to estimate the mean production at different ages (prediction) and the second (Case 2) was to estimate and extrapolate the mean production in previously set ages (six and seven years old - the usual regulatory rotation age in the Minas Gerais State).

Besides the regional categorical variables, namely: spacing, genotype, cutting cycle and predominant soil in Case 1, the variables in Tables 1 and 2 were also taken into consideration. Production was the output variable (m³/ha). Training and validation percentages were: 100(0), 50(50), 40(60), 30(70), 20(80) and 10(90). The numbers in brackets refer to the percentage given to the data in order to validate the trained neural networks. The trained network database just presented variation in this level.

The aim of Case 2 was to assess prediction efficiency by using the climatic and edaphic variables, and to assess the age generation of the mean productivity map for the entire state. The applied categorical output variables were the sub-region, spacing, genotype, cutting cycle, predominant soil, altitude; highest, lowest and mean temperatures, and minimum, maximum and the mean precipitation during the sampled period (Table 2). Six and seven-years-old productions (m³/ha) were chosen as output variables. Only the 7 years old parcels (IMA7)- acceptable variation between 78.1 and 90 months, and 6 years old parcels (IMA6)- acceptable variation between 66 and 78 months, were selected in the IFC database, since the used database (Table 2) did not contain parcels from all the cities in the herein studied state. Information on temperature, precipitation and climate (Köppen-Geiger classification), extracted from http://www.ipef.br/geodatabase were also employed, besides the variables in Table 2. The percentage of data for training and validation in case 2 were 100(0), 95(5), 90(10) and 85(15).

### Training and generalizing artificial neural networks

The network training was performed in the parcel level because the output variable “volume”, presented variation in this level in both cases. However, the validation was performed in the block level because the edaphic and climatic variables used in the training network database just presented variation in this level. The Neuroforest was the software used to train and apply the networks. The trained network was the Multilayer Perceptron (MLP), with three layers. The Resilient propagation (RPROP) (Riedmiller and Braun, 1993) was the employed algorithm, since it adapts its weight of each step according to the local gradient information. Such adaptation is not influenced by the behavior of the gradient.

The amount of neuron in the entrance layer fluctuated according to the number of variables taken into consideration in each study. Twelve neurons were used in the intermediate layer and one, in the

### Table 1. Amplitude, mean values and standard deviation for age, dominant height (Dh), basal area (B) and volume in parcel level, in the sampled area of the state of Minas Gerais, Brazil.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (months)</td>
<td>11.97</td>
<td>357.34</td>
<td>61.17</td>
<td>35.18</td>
</tr>
<tr>
<td>Dh (m)</td>
<td>6.00</td>
<td>59.83</td>
<td>24.44</td>
<td>6.59</td>
</tr>
<tr>
<td>N (m³/ha)</td>
<td>0.70</td>
<td>58.91</td>
<td>18.37</td>
<td>6.92</td>
</tr>
<tr>
<td>Yield (m³/ha)</td>
<td>3.80</td>
<td>1158.42</td>
<td>197.16</td>
<td>118.75</td>
</tr>
</tbody>
</table>

### Table 2. Amplitude and standard deviation of the edaphic and climatic variables used as input variables in train neural networks to estimate eucalyptus development, production and productivity in the state of Minas Gerais, Brazil.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Year</th>
<th>Minimum</th>
<th>Mean</th>
<th>Maximum</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature (°C)</td>
<td>14.69</td>
<td>19.85</td>
<td>29.69</td>
<td>3.82</td>
<td></td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td></td>
<td>72.77</td>
<td>77.83</td>
<td>81.83</td>
<td>2.74</td>
</tr>
<tr>
<td></td>
<td>2007</td>
<td>20.08</td>
<td>69.68</td>
<td>114.91</td>
<td>20.50</td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>19.86</td>
<td>107.23</td>
<td>162.56</td>
<td>30.10</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>19.98</td>
<td>115.68</td>
<td>160.86</td>
<td>37.62</td>
</tr>
<tr>
<td>Annual precipitation averages (mm)</td>
<td>2010</td>
<td>72.77</td>
<td>77.83</td>
<td>81.83</td>
<td>2.74</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>18.52</td>
<td>115.73</td>
<td>179.84</td>
<td>38.68</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>19.78</td>
<td>83.01</td>
<td>109.55</td>
<td>23.10</td>
</tr>
<tr>
<td></td>
<td>2013</td>
<td>19.32</td>
<td>122.26</td>
<td>190.50</td>
<td>40.99</td>
</tr>
<tr>
<td>Wind speed (m/s)</td>
<td>1.27</td>
<td>2.97</td>
<td>4.30</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>Total radiation (MJ/m²/day)</td>
<td>12.88</td>
<td>14.81</td>
<td>17.23</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>31.020,50</td>
<td>32.979,50</td>
<td>34.721,94</td>
<td>1.256,72</td>
<td></td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td>8.76</td>
<td>14.92</td>
<td>24.57</td>
<td>5.57</td>
<td></td>
</tr>
<tr>
<td>Annual precipitation averages (mm)</td>
<td>4.10</td>
<td>5.98</td>
<td>8.41</td>
<td>1.45</td>
<td></td>
</tr>
</tbody>
</table>
The correlation between the observed and estimated \( r_{Y\hat{Y}} \) volumes of eucalyptus stands in Case 1.

<table>
<thead>
<tr>
<th>Percentage for training</th>
<th>Percentage for validation</th>
<th>Status</th>
<th>RMSE (%)</th>
<th>( r_{Y\hat{Y}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0</td>
<td>Training</td>
<td>5.15</td>
<td>0.9929</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td>Training</td>
<td>4.65</td>
<td>0.9941</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td>Validation</td>
<td>5.13</td>
<td>0.9931</td>
</tr>
<tr>
<td>40</td>
<td>60</td>
<td>Training</td>
<td>4.64</td>
<td>0.9943</td>
</tr>
<tr>
<td>40</td>
<td>60</td>
<td>Validation</td>
<td>4.97</td>
<td>0.9934</td>
</tr>
<tr>
<td>30</td>
<td>70</td>
<td>Training</td>
<td>4.65</td>
<td>0.9943</td>
</tr>
<tr>
<td>30</td>
<td>70</td>
<td>Validation</td>
<td>4.77</td>
<td>0.9939</td>
</tr>
<tr>
<td>20</td>
<td>80</td>
<td>Training</td>
<td>4.64</td>
<td>0.9943</td>
</tr>
<tr>
<td>20</td>
<td>80</td>
<td>Validation</td>
<td>5.65</td>
<td>0.9915</td>
</tr>
<tr>
<td>10</td>
<td>90</td>
<td>Training</td>
<td>4.65</td>
<td>0.9943</td>
</tr>
<tr>
<td>10</td>
<td>90</td>
<td>Validation</td>
<td>6.17</td>
<td>0.9898</td>
</tr>
</tbody>
</table>

Assessing the artificial neural network forecast

The assessment of the artificial neural network forecast in training and validation stages were done through statistics and graphic analysis. The employed statistics were the mean absolute differences (\( \bar{e} \)), the correlation between estimated and observed volumes (\( R_{Y\hat{Y}} \)), root mean square error (RMSE) and the relative percentage error (RE %).

\[
\bar{e} = n^{-1} \sum_{i=1}^{n} |Y_i - \hat{Y}_i|, \quad \text{RMSE} = \sqrt{n^{-1} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}, \quad \text{Bias} = n^{-1} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)
\]

\[
R_{Y\hat{Y}} = \frac{n^{-1} \left( \sum_{i=1}^{n} (Y_i - \hat{Y}_m)(\hat{Y}_i - \bar{Y}) \right)}{\sqrt{\left( n^{-1} \sum_{i=1}^{n} (Y_i - \hat{Y}_m)^2 \right) \left( n^{-1} \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 \right)}},
\]

\[
\hat{Y}_m = n^{-1} \sum_{i=1}^{n} \hat{Y}_i, \quad \text{ER}\% = 100n^{-1} \sum_{i=1}^{n} \frac{Y_i - \hat{Y}_i}{Y_i}
\]

Wherein: \( Y_i, \hat{Y}_i \) and \( \bar{Y} \) are the observed value, the model estimation value and mean observed values, respectively, and “n” is the number of cases.

RESULTS

Case 1

Table 3 shows root mean square error (RMSE) estimations and the correlation between observed and estimated volumes \( r_{Y\hat{Y}} \), according to the training and validation percentage, in Case 1. Figures 1 and 2 show the dispersion graphs of observed and estimated volumes; and the corresponding histograms of residue distribution frequency, which used 100, 50, 40, 30, 20 and 10% of the data kept for network training in Case 1.

Case 2

Table 4 presents root mean square error (RMSE %) estimations and the correlation between observed and estimated values \( r_{Y\hat{Y}} \), according to the validation and training percentage. Figures 3 and 4 introduce graphs of estimated and observed volumes and histograms that correspond to the residue distribution frequency, which used 100, 95, 90 and 85% of the data kept for network training. Figures 5 and 6 present the productivity maps for Minas Gerais at the ages of 7 (IMA7) and 6 (IMA6) using the neural network from Case 2, which was trained with 100% of the data.

DISCUSSION

The link between the observed and the estimated volumes (Tables 3 and 4) indicates the strength and direction between the two variables. The closest to 1 it is, the greater the correlation between the variables. The root mean square error assesses the error between the observed and the estimated volumes; the greater the RMSE is, greater the accuracy (Mehtätalo et al., 2006). When the number of observations (number of parcels and blocks in the present study) is relatively large, the RMSE forecast presented in Tables 3 and 4 may be understood as residual standard error.

The statistics obtained in the validation predictions were lower than 6% in the network training with more than 20% of data available; and equals to 6.17% in the
training network with 10% of the data availability. The correlation between the observed and the estimated productions was 0.99 in all the cases (Table 3). This accuracy is adequate for prediction in the parcel level. Besides, it is possible to observe that the errors often follow the normal distribution and the relative errors (RE %) float around the mean 0. The graphic analysis of the errors (residues) was also employed to assess the neural network models and it contained histograms presenting the frequency of the cases through relative error category percentage and cross-validation graph (observed volume versus estimated volume). The obtained relative error distributions met the results often obtained in production and yield models used in Brazil.

According to Tables 3 and 4, the RMSE and the correlation between the observed and the estimated values remained constant in trainings that had used 100 to 10% of the data in this stage. However, in Case 2 in which validation did not include the stands variable “inventory”, the RMSE forecasts were almost 50% of those observed in the network training. Besides, the validation error estimates were satisfactory. The neural network in Case 2 must be trained with 100% of the data available in order to extrapolate the whole state productivity; so that the relative error margin (RE %) is 10.36% with ryy correlation of 0.88.

More than 90% of the errors were close to 7.5%, which is an excellent result for parcel-level data. The range of errors (RE %) is wider although acceptable for the reforestation investment analysis, since Case 2 just employs edaphic and climatic variables. All the cases (Figures 1 to 4) presented normal error distribution, and
Case 1 presented a more leptokurtic shape. It is still possible to observe relatively high correlations between the observed and the estimated values, over 80% (Table 4), despite the accuracy loss due to the adoption of the edaphic and climatic variables without the IFC data. The error diffusion (Figures 3 and 4) changes in this case, that is, the residual variance may be directly interpreted through the RMSE, due to the large number of observations. Therefore, in this case, it is an excellent approximation to the residual standard error. The error distribution in Case 2 was normal in several projection ranges; it presented 90% of errors between 25 and -25% variance. Neural networks in Case 2 may be employed to the new investments in the eucalyptus plantation of the states with a good margin of error, given that the residual graphic analysis was performed in the plot level and that the projection took into account different age ranges in the absence of IFC data.

Figure 2. Observed versus estimated volumes and histograms corresponding to the percentage error frequency using 30, 20 and 10% of the data to train artificial neutral networks in Case 1.
Table 4. Root mean square error (RMSE %) and the correlation between the observed and estimated (r��) volumes of eucalyptus stands in Case 2.

<table>
<thead>
<tr>
<th>Percentage of training data</th>
<th>Percentage of validation data</th>
<th>Status</th>
<th>RMSE (%)</th>
<th>(r��)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0</td>
<td>Training</td>
<td>10.36</td>
<td>0.8830</td>
</tr>
<tr>
<td>95</td>
<td>5</td>
<td>Training</td>
<td>12.18</td>
<td>0.8377</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validation</td>
<td>21.83</td>
<td>0.5295</td>
</tr>
<tr>
<td>90</td>
<td>10</td>
<td>Training</td>
<td>11.63</td>
<td>0.8555</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validation</td>
<td>22.14</td>
<td>0.5387</td>
</tr>
<tr>
<td>85</td>
<td>15</td>
<td>Training</td>
<td>11.21</td>
<td>0.8640</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validation</td>
<td>22.22</td>
<td>0.5355</td>
</tr>
</tbody>
</table>

Figure 3. Observed versus estimated volumes and the histograms corresponding to percentage error frequency using 100% and 95% of the data to train artificial neural networks in Case 2.

The IMA7 map (Figure 5) presents more productivity classes than the IMA6 map (Figure 6), and it is justified by the different results of growth curves of clones in different regions of the state. The curves were more similar until the age of six, but from this point on, they differ from each other and show different productivity and yield rates. The maps generated in the present study can be employed in studies on the expansion of cultivation areas. However, the overlapping of an updated land use map is necessary to enable this project’s implementation. Future studies are necessary to map the grazing lands in the state, especially the degraded ones; as well as to assess the potential timber production in the areas cultivated in an agroforestry system layout in the state. Between 30 and 40% of training basis is enough for production modeling and for prediction purposes in IFC databases or for the IFC added to physiographic, edaphic and climatic information. Such percentage presents great
coverage of eucalyptus plantation areas in Minas Gerais State. All the data available must be used to map the productivity classes (Case 2); unless there is too much data for the combinations, and for the categorical and continuous variables. In that case, the neural network itself may be used to decide what variables and data are relevant for the modeling or for the analysis; or the main component analysis may be previously used to reduce the database without causing representative losses to it. Besides, a specific neural network may be developed and applied at the time to decide data which are really important for the modeling and mapping processes.

The productivity simulated by Borges (2012) shows the potential that may be reached if adequate silvicultural practices and genotypes are employed in the area, without limitations caused by physiographic-nutritional features. The author observed that productivity rises as precipitation increases and maximum temperature diminishes. The mean production varied from 42.3 to 73.4 m$^3$/ha/year in six years old trees. The author concluded that precipitation, solar radiation, rain distribution throughout the years, and maximum temperature influence the potential productivity. The writer estimated potential productivities between 40 and 60 m$^3$ ha$^{-1}$ ano$^{-1}$ at seven-years-old trees in the state of Minas Gerais; whereas Guimarães et al. (2007) estimated potential productivities that float from 6 to 23 and from 10 to 50 m$^3$ ha$^{-1}$year$^{-1}$, for high and medium technological levels, respectively.

Borges (2012) and Guimarães et al. (2007) employed...
Figure 5. Productivity at the age of 7 (IMA7) using an artificial neural network in Case 2.
the procedural model. The present study estimated production of 20 a 50 m³ ha⁻¹ year⁻¹ for the ages of 6 and 7 by applying the RNA using data from forest surveys conducted in the last ten years. Those forecasts reflect the climatic and silvicultural reality in the stands plantation up to 2014.

The results in the present paper show that the mean production of eucalyptus stands in the state of Minas
Gerais present volumes between 35 and 40 m$^3$ ha$^{-1}$ for seven years old trees. The generated maps help defining the reforestation public policies in the state of Minas Gerais. They also help subsidizing the expansion plans for the plantation areas in the state. The configuration and topology herein defined are the starting points to the development of artificial neural network models to be used in other Brazilian states and in other countries; and also to the mapping of productivity classes. Differently from the traditional statistical approaches demand of the fulfillment of statistical assumptions, the neural networks do not demand such requirement (Jensen et al., 1999) and they may be employed in modeling by using stands variables and biophysical parameters.

The aim of the present study was not to simulate climate scenario, but to demonstrate the neural network efficiency predictions; and to use them in mapping the productivity classes in the state of Minas Gerais. The used setup may be a starting point for similar studies focused on mapping carbon productivity or storage in aerial or soil biomass. The continuous variables and the categories employed in the present study are enough to estimate the future production in eucalyptus stands.

Besides, RNA may also be employed to predict production improvement and to project forthcoming climate conditions (Ashraf et al., 2015). Thereby, it is possible to estimate upcoming productions in different climate scenarios. Such estimations are important due to climate changes from the last years and to the neural networks efficiency in this type of modeling.

**CONFLICT OF INTERESTS**

The authors have not declared any conflict of interests.

**REFERENCES**


Borges JS (2012). Modulador edáfico para uso em modelo ecosiológico e produtividade potencial de povoamentos de eucalipto. Thesis (PhD in Forest Science) - Universidade Federal de Viçosa, Viçosa – Brasil, MG. P 70.


Trevizol Júnior TL (1985). *Análise de um modelo compatível de crescimento e produção em plantações de Eucalyptus grandis (W.Hil ex-maiden)*. Dissertation (Master in Forest Science) - Universidade Federal de Viçosa, Viçosa – Brazil, MG, P 74.
