Tree biomass estimation in regenerating areas of tropical dry vegetation in northeast Brazil

Everardo Sampaio a,1, Peter Gasson b,*, Amelia Baracat b, David Cutler b, Frans Pareyn c,**, Kleber Costa Lima c,2

a Departamento de Energia Nuclear, Universidade Federal de Pernambuco, Av. Prof. Moraes Rego, no. 1235, Cidade Universitária, Recife, PE, Brazil
b Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, United Kingdom
c Associação Plantas do Nordeste - APNE, Rua Dr. Nina Rodrigues, 265, Iputinga, 50.731-280 Recife, PE, Brazil

1. Introduction

Allometric equations for the estimation of biomass have been developed for a wide range of species (Ter-Mikaelian and Korzukhin, 1997; Zianis and Mencuccini, 2004) and types of vegetation (Haase and Haase, 1995; Sampaio and Silva, 2005; Chave et al., 2005; Keller et al., 2001; Ketterings et al., 2001; Brown, 2002; Zianis and Mencuccini, 2004; Nogueira et al., 2008). The species diversity is lower in dry tropical regions than in the humid tropics and fewer equations have been developed (Chidumayo, 1997; Bellefontaine et al., 2000; Sampaio and Silva, 2005). In some cases they are for single species but there are a few for groups of species (Cairns et al., 2003; Sampaio and Silva, 2005). Although tropical dry forests are the most extensive land cover type in the tropics, fewer studies have been conducted to estimate their biomass and carbon stock (Urquiza-Haas et al., 2007). They depend on valid allometric equations.

The relatively abundant literature on the development of equations and on their theoretical aspects (Ter-Mikaelian and Korzukhin, 1997; Ketterings et al., 2001; Enquist, 2002; Zianis and Mencuccini, 2004; Pilli et al., 2006) contrasts with the scarcity of studies validating them in the field (Chave et al., 2005; Nogueira et al., 2008), which are essential before such equations can be reliably used in the estimation of biomass of large areas. This necessity is reinforced by the fact that many of these equations have been based on few plants (Brown et al., 1995; Ketterings et al., 2001), rarely more than a few dozens, and cover a large range of regions with a wealth of low dominance species, equations have usually been developed for groups of species in a given area (Overman et al., 1994; Keller et al., 2001; Ketterings et al., 2001; Brown, 2002; Zianis and Mencuccini, 2004; Chave et al., 2005; Nogueira et al., 2008). The species diversity is lower in dry tropical regions than in the humid tropics and fewer equations have been developed (Chidumayo, 1997; Bellefontaine et al., 2000; Sampaio and Silva, 2005). In some cases they are for single species but there are a few for groups of species (Cairns et al., 2003; Sampaio and Silva, 2005). Although tropical dry forests are the most extensive land cover type in the tropics, fewer studies have been conducted to estimate their biomass and carbon stock (Urquiza-Haas et al., 2007). They depend on valid allometric equations.

The relatively abundant literature on the development of equations and on their theoretical aspects (Ter-Mikaelian and Korzukhin, 1997; Ketterings et al., 2001; Enquist, 2002; Zianis and Mencuccini, 2004; Pilli et al., 2006) contrasts with the scarcity of studies validating them in the field (Chave et al., 2005; Nogueira et al., 2008), which are essential before such equations can be reliably used in the estimation of biomass of large areas. This necessity is reinforced by the fact that many of these equations have been based on few plants (Brown et al., 1995; Ketterings et al., 2001), rarely more than a few dozens, and cover a large range of
plant sizes (Brown et al., 1995), so that each size group is represented by only a few individuals (Fehrmann and Kleinn, 2006). In the dry tropics, the size range is more limited than in humid tropical and temperate forests and some studies have used a relatively large number of plants. Cairns et al. (2003), for instance, used 698 plants for the vegetation on the Mexican Yucatan Peninsula. A modified version of the Yucatan equation was later applied to a large set of plot data to estimate forest biomass in the peninsula (Urquiza-Haas et al., 2007). In the Brazilian caatinga, Sampaio and Silva (2005) developed equations for individual species and also for groups of species, using a relatively large set of plants: 30 for each of nine species. However, no validation of allometric equations for dry tropical species could be found in the literature of the last decade. In the humid Amazon forest, equations developed for the central region overestimated the biomass of about 300 plants of the southern region (Nogueira et al., 2008).

The semi-arid region of NE Brazil has approximately one million square km originally covered by a shrub and tree vegetation called caatinga. About half of this area is still covered by native vegetation that is used for the production of fuel wood and slash and burn agriculture, resulting in a mosaic of vegetation with different regeneration periods (Sampaio, 1995). Mature preserved caatinga areas are small and few in number. The regeneration areas are dominated by relatively few species and Sampaio and Silva (2005) developed aboveground biomass equations for the most important ones. Equations for their fuel wood biomass and volume have also been published (PNUD/FAO/IBAMA, 1992). However, the biomass equations were based on trees growing in mature areas, to cover a larger range of sizes, and have not been validated in field conditions, particularly in regenerating locations where the vegetation is cut every 10–15 years. At this growth stage, the range of sizes is narrower than in the mature areas and the range of trees cut for fuel wood (above 2–3 cm diameter at the base of the stem) still narrower. The architecture of plants in regenerating areas may also be different, since most of them coppice and the number of sprouts is progressively reduced as the plants grow taller (Sampaio et al., 1993, 1998).

This study was aimed at validating the equations already developed for the caatinga species and proposes the development of new ones specific for caatinga vegetation regenerating after being cut for firewood.

2. Materials and methods

The work was conducted, in 2002, in three localities in the caatinga of Pernambuco, two in the municipality of Sertânia and one in the municipality of Serra Talhada. In Sertânia, we chose one area of native vegetation of circa 40 ha, at the IPA Experimental Station (08°04'02.7"S and 37°12'33.1"W) and another privately owned site of circa 20 ha in the district of Carolina (08°15'33.0"S and 37°32'52.1"W). In Serra Talhada, we chose an area of 80 ha, at the IPA Experimental Station (07°55'46.4"S and 38°17'20.0"W). The rainfall in these two municipalities is concentrated between January and May, with total annual means of 483 mm in Sertânia, and 626 mm in Serra Talhada, and high variation from year to year (Figueirôa et al., 2006). The mean annual temperatures vary little, ranging between 26 and 28 °C. The three localities have soils with a coarse sandy texture and medium fertility (Figueirôa et al., 2006).

In Serra Talhada and in Sertânia, at the IPA Experimental Station, we marked, both in the wet season and in the dry season, at each locality, 60 plants of each of the four species most used for production of fuel wood: catingueira, Coesalpinia pyramidalis Tul.; jurema preta, Mimoso tenuiflora (Wild.) Poir.; mameleiro, Chorisia sonderianus Muell. Arg.; and jurema de imbirica, Mimosa ophthal- mocentra Mart. ex Benth. In Carolina, C. sonderianus and M. ophthalmostilico were absent and the vegetation was dominated by C. pyramidalis and M. tenuiflora, so we cut 60 trees of each of the two species present. The trees, chosen at random, were at least 2 cm in diameter at breast height, measured 1.3 m aboveground level (DBH), which is the size cut for fuel wood in the region. For each tree, the circumferences at the stem base (CSB) and at breast height (CBH), total height (H) and the two major orthogonal axes of the crown projection area were measured. Each tree was then cut at the base and separated into pieces (stem, large and small branches, twigs and leaves) that were weighed in the field and sampled to determine the dry weight in the laboratory. To determine wood density (WD) we took discs 2.5 cm thick, from the trunk and largest branches, each metre up the plant, and the disc volumes were measured by immersion, and their weight determined after drying at 103 °C (Zakia et al., 1992).

At each locality, we collected fertile branches of the four species for confirmation of their identity. This voucher material is deposited at the Herbário Dândaro de Andrade Lima (IPA).

We developed allometric equations considering the dry weight of each plant as the dependent variable, and height, diameter at the stem base (DSB), diameter at breast height (DBH), crown area (CA) and wood density as independent variables, taken individually or combined in groups of two, three or four variables, used as single variables or as multiple variables. DSB and DBH were calculated assuming that the stems were circular and the area of the crown was calculated assuming an ellipse. The data were adjusted for different types of equation models (linear, quadratic, power, exponential and logarithmic), found in the literature, using the individual or combined variables. We developed equations for increasing sets of data, from the 60 plants of each species, at each season (wet or dry) and locality, to the total of all 1200 plants. The equations (model and parameter values) for one season and for each species and site were validated against the sets of data for the other season, using their real input variables (H, DSB, DBH, CA and WD) and comparing predicted and real biomasses (Nelson et al., 1999). If they were considered as valid, the data of the two seasons were pooled together, new equations developed and validated against the sets of data for the same species in the other sites. If they were also considered as valid, a single new equation was developed for each species. These equations were validated against the small set of data for the same species collected by Sampaio and Silva (2005) but were not validated using data from this present work because all of them had been used to develop the equation. These equations were also used to estimate the biomass of the other species, in the search for equations of more general use but these comparisons cannot be taken as validation. Fitness of the equations was compared by their regression coefficients ($R^2$), mean squares of residuals (MSR) or square root of MSR (SMSR), sum of squares of the differences between predicted and real values (PRESS) or error mean square (EMS) and mean difference between predicted and real values or aggregate difference (AD). The most important validation criterion was the sum of differences between predicted and real values in relation to sum of real values or AD in relation to the average real value (relative AD). In the estimates of the log-transformed equations we applied the correction factor (CF), described in the literature (Chave et al., 2005). If the relative AD was above 0.2 (20% deviation from the real value) the equation was considered as not valid, regardless of the other criteria; if it was below 0.2, the other criteria were considered. When comparing different equation models for the same data set, the best model was selected by ranking them according to the goodness of each criterion and summing the rank positions (Salas, 2002). The Fournival index was used to correct the SMSR of the transformed models (Segura and Andrade, 2008). When validating an equation against another data set, relative AD and EMS were used. We also compared our biomass data with estimates made with equations and their parameters developed
for other vegetation types, published in the literature (Overman et al., 1994; Haase and Haase, 1995; Ter-Mikaelian and Korzukhin, 1997; Kumar et al., 1998; Nelson et al., 1999; Chambers et al., 2001; Keller et al., 2001; Cairns et al., 2003; Zanis and Mencuccini, 2004; Chave et al., 2005), using our real input variables. Due to the large number of comparisons and validation tests, only the most significant results are presented.

3. Results and discussion

The trees of the four species varied considerably in size as shown in Table 1. As expected, the linear measurements have a narrower range of variation than those of bi- and tridimensional measurements. Tree height (H) varied from 2 to 10 m (a 5-fold range) and the diameter at breast height (DBH) from 2 to 34 cm (17 times), crown area (CA) varied from 0.3 to 107 m² (356 times) and biomass (B) from 1 to 392 kg (392 times). In general, M. ophthalmocentra and M. tenuiflora trees were largest, C. sonderianus the smallest and C. pyramidalis intermediate in size. The majority of the plants were concentrated in a narrow size range. For example, more than 90% of the plants of all species had heights between 4 and 7 m and more than 65% of the plants of M. ophthalmocentra, M. tenuiflora and C. pyramidalis had DBH between 5 and 15 cm while 98% of those of C. sonderianus were less than 10 cm DBH (Table 2). Wood density varied little within each species. The mean densities of C. sonderianus (0.88 g cm⁻³) and C. pyramidalis (1.02 g cm⁻³) were a little higher (t-test p < 0.01) than those recorded (0.77 and 0.92 g cm⁻³) from mature areas (Sampaio and Silva, 2005).

A wide range of models and equations was well adjusted (significant R² at p < 0.01) for our data for each species, locality and season of cutting, and their various combinations. Equations for trees harvested in the rainy and dry seasons were not significantly different and were valid for the set of the other season. In general, the same occurred with equations for the different localities, with the unique exception of M. tenuiflora which had trees with higher biomasses for the same DBH values in Serra Talhada than in the two locations in Sertânia. Therefore, with the exception of M. tenuiflora, the trees from all the sites were combined for further analysis and comparisons.

The regression coefficients indicate that relatively high proportions of variability in tree mass can be explained by the stem diameter, height and crown area. Considering each variable separately, the best fit was generally with DBH (Table 3), with little difference comparing to DBH². Height, crown area and particularly wood density, as separate independent variables, resulted in worst fit, frequently with non-significant regression coefficients. Multiplying or adding height, crown area or wood density to DBH, sometimes resulted in better fittings but the improvements were small. As has been concluded in other work, circumference or diameter are the best indicators of tree mass (Ter-Mikaelian and Korzukhin, 1997). Accurate height measurements are difficult to make and seldom available in large surveys and height is usually correlated with DBH, therefore many authors suggest not to include it in biomass estimations (Haase and Haase, 1995; Brown, 2002; Chave et al., 2005; Smith and Whelan, 2006; Nogueira et al., 2008). In our case, several significant correlations between H and DBH indicated that it was not valid to use them as separate independent variables and explains the small improvement in the fitness when using DBH × H or DBH² × H. Crown area is even more difficult to measure than height, it has seldom been used in equations and its contribution tends to be small (Samba et al., 2001; Sampaio and Silva, 2005). In some of our cases, CA was correlated with H and/or DBH and its inclusion did not improve the fitness above those obtained with DBH or DBH × H. Wood density is also difficult to determine and its contribution to equations developed for single species is small because it varies little among individuals of the same species. However, species average density and average or regional stand height may be important variables in equations developed for sets of many different species to be applied at different sites (Chambers et al., 2001; Chave et al., 2005; Fehrman and Kleinn, 2006). Density has frequently been related to the scaling coefficient of the power equation (Ketterings et al., 2001; Pilli et al., 2006).

In several cases, the best fit obtained with one type of model (linear, for example) was just a little superior to those encountered with other types (quadratic or power, for example) and both were significant. In general, the narrower the variation of biomass of the plants considered, the higher the similarity among types of models. This happened mainly with C. sonderianus, the species with the narrowest range (Tables 2 and 3). In the vast majority of cases, the

---

**Table 1**

<table>
<thead>
<tr>
<th>Species</th>
<th>DBH (cm)</th>
<th>Height (m)</th>
<th>Crown area (m²)</th>
<th>Biomass (kg)</th>
<th>Density (g cm⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. ophthalmocentra</td>
<td>24</td>
<td>9</td>
<td>1.3</td>
<td>1</td>
<td>392</td>
</tr>
<tr>
<td>M. tenuiflora</td>
<td>34</td>
<td>10</td>
<td>107</td>
<td>4</td>
<td>240</td>
</tr>
<tr>
<td>Caesalpinia pyramidalis</td>
<td>19</td>
<td>10</td>
<td>2.1</td>
<td>3</td>
<td>129</td>
</tr>
<tr>
<td>Croton sonderianus</td>
<td>13</td>
<td>8</td>
<td>0.3</td>
<td>1</td>
<td>44</td>
</tr>
</tbody>
</table>

* Minimum values of 2 cm and 2 m.

**Table 2**

<table>
<thead>
<tr>
<th>Species</th>
<th>DBH (cm)</th>
<th>Height (m)</th>
<th>Crown area (m²)</th>
<th>Biomass (kg)</th>
<th>Density (g cm⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. tenuiflora 1</td>
<td>8</td>
<td>42</td>
<td>94</td>
<td>51</td>
<td>4</td>
</tr>
<tr>
<td>M. tenuiflora 2</td>
<td>6</td>
<td>66</td>
<td>34</td>
<td>14</td>
<td>–</td>
</tr>
<tr>
<td>M. ophthalmocentra</td>
<td>13</td>
<td>121</td>
<td>73</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td>Caesalpinia pyramidalis</td>
<td>15</td>
<td>180</td>
<td>135</td>
<td>30</td>
<td>–</td>
</tr>
<tr>
<td>Croton sonderianus</td>
<td>109</td>
<td>126</td>
<td>5</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

**Table 3**

<table>
<thead>
<tr>
<th>Species</th>
<th>DBH (cm)</th>
<th>R²</th>
<th>DBH × H (cm × m)</th>
<th>R²</th>
<th>DBH × C (cm × m²)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. ophthalmocentra</td>
<td>0.4369D1.9609 0.81</td>
<td>0.0893DH1.4582 0.73</td>
<td>0.162DC+ 9.12 0.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. pyramidalis</td>
<td>0.3129D1.8838 0.68</td>
<td>0.0934DH1.3904 0.69</td>
<td>0.145DC+ 8.17 0.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. sonderianus</td>
<td>0.4171D1.5601 0.62</td>
<td>0.0867DH1.2852 0.67</td>
<td>0.140DC+ 2.60 0.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. tenuiflora 1</td>
<td>0.3344D1.9048 0.70</td>
<td>0.071ODH1.5120 0.69</td>
<td>0.129DC+ 9.35 0.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. tenuiflora 2</td>
<td>0.4138D1.7718 0.64</td>
<td>0.0798DH1.4501 0.66</td>
<td>0.123DC+ 8.79 0.79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
power equations showed a better fit to the data. The vast majority of studies, with very diverse types of vegetation, have arrived at the same conclusion (Chave et al., 2005; Pilli et al., 2006) and the power equation is backed by a large body of theory on plant growth, structure and scaling (Enquist, 2002; Niklas, 2004; Zianis and Mencuccini, 2004; Fehrmann and Kleinn, 2006).

The masses estimated from the power equations were very close to the genuine masses of the set of trees of each species, in all cases with differences of less than 10% of the total masses (relative aggregate difference). This result reflects that the fit of the equations was good. For individual trees, the estimates sometimes resulted in much higher or lower values than the true mass (higher PRESS or EMS) but the deviations tended to be compensated by those of other plants. In general, deviations were greater for the larger plants (illustrated for the species with lowest and highest deviations, Figs. 1 and 2), which were also represented by fewer sampled individuals (Table 2).

On the other hand, using the equations previously developed for the caatinga species (Sampaio and Silva, 2005) resulted in estimates much higher than the real masses: at least 15% for M. ophthalmoconcentra and up to 80% for C. sonderianus. This overestimation occurred with the equations specifically developed for M. tenuiflora, C. pyramidalis and C. sonderianus (30 plants each) and also with the general caatinga equation, based on 270 plants (Sampaio and Silva, 2005). Therefore, equations developed for trees in mature vegetation are not applicable to trees in regenerating sites, usually cut for firewood. In general, equations developed for other tropical vegetations (Overman et al., 1994; Haase and Haase, 1995; Kumar et al., 1998; Nelson et al., 1999; Chambers et al., 2001; Keller et al., 2001; Zianis and Mencuccini, 2004; Fehrmann and Kleinn, 2006) also overestimated the real masses of the trees of the regenerating sites. The closest estimate, even more than using the mature caatinga equations, was obtained with Cairns et al. (2003) equation for large plants of the Mexican Yucatan, which has a physiognomy similar to that of caatinga. The differences in estimates with the equations for mature and regenerating sites stress the importance of validating the equations before applying them under different conditions, even if these seem to be small.

The differences between the estimates using the equations for mature and for regenerating sites reflect the fact that, for the same stem diameters, the trees in the regenerating sites had less biomass (Fig. 3). The differences in height only partially explain the differences in biomass because the estimates using DBH \( \times \) H were still higher with the equation for the mature sites (Fig. 4). With the

---

**Fig. 1.** Aggregate differences (ADs) and mean quadratic differences (MQE) of the real biomasses of plants in classes of increasing stem diameter compared to biomasses estimated by equations based on diameter at breast height, in relation to the average biomasses in these classes. Lines labelled 1 refer to the equation developed in this paper (Table 3) and lines labelled 2 to the general equation developed by Sampaio and Silva (2005).

**Fig. 2.** Aggregate differences (ADs) and mean quadratic differences (MQE) of the real biomasses of plants in classes of increasing stem diameter compared to biomasses estimated by equations based on diameter at breast height, in relation to the average biomasses in these classes. Lines labelled 1 refer to the equation developed in this paper (Table 3) and lines labelled 2 to the general equation developed by Sampaio and Silva (2005).

**Fig. 3.** Biomass (kg) estimated from diameter of stem at breast height (DBH, cm), compared with estimates made from equations in the literature (Z&M 2004, Zianis and Mencuccini, 2004; S&S 2005, Sampaio and Silva, 2005). M. tenuiflora 1 = from Serra Talhada; M. tenuiflora 2 = from Sertãozinho.

**Fig. 4.** Biomass (kg) estimated from the diameter at breast height multiplied by tree height (DBH in cm × H in m), compared with estimates made from equations in the literature (S&S 2005, Sampaio and Silva, 2005; C et al. 2003, Cairns et al., 2003). M. tenuiflora 1 = from Serra Talhada; M. tenuiflora 2 = from Sertãozinho.
equation by Cairns et al. (2003), using $DBH^2 \times H$, the differences were slightly reduced, indicating that the form of caatinga trees is similar to those in the Yucatan, which are intermediate between the mature and the regenerating caatinga. Our results and those from the Yucatan were based on measurements from hundreds of plants, chosen to represent a definite vegetation pattern and not to represent a regular size increase along the species size range. On the other hand, the similarity between the estimates with the regenerating caatinga and the Yucatan equations reinforce the need to search for equations that are more general for similar vegetation types (Enquist, 2002; Niklas, 2004; Zianis and Fehrmann, 2006).

Since the biomasses of trees, with the same DBH and height, were lower in the regenerating sites, the difference is in the branches, which are fewer, shorter and/or thinner in regenerating sites. This hypothesis is reinforced by the estimates of crown areas, which are also smaller with equations for regeneration trees (Table 4) than with the equations of trees from mature caatinga. It is possible that in the mature areas each tree has more space in which to develop a large crown. Also, the repeated cutting in the central Amazon. Forest Ecology and Management 117, 149–167.


2005), M. tenuiflora 1 = from Serra Talhada; M. tenuiflora 2 = from Sertânia.

**Table 4**

<table>
<thead>
<tr>
<th>Species</th>
<th>DBH (cm)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. ophthalmocentra</td>
<td>0.591D$^{1.341}$</td>
<td>0.64</td>
</tr>
<tr>
<td>C. pyramidalis</td>
<td>0.650D$^{1.234}$</td>
<td>0.48</td>
</tr>
<tr>
<td>C. sonderianus</td>
<td>0.394D$^{1.362}$</td>
<td>0.41</td>
</tr>
<tr>
<td>M. tenuiflora 1</td>
<td>0.958D$^{1.231}$</td>
<td>0.47</td>
</tr>
<tr>
<td>M. tenuiflora 2</td>
<td>0.871D$^{2.007}$</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Fig. 5. Crown area (m$^2$) estimated from diameter at breast height (DBH, cm), compared with estimates made from equations in the literature (S & S 2005, Sampaio and Silva, 2005). M. tenuiflora 1 = from Serra Talhada; M. tenuiflora 2 = from Sertânia.

**4. Conclusions**

A group of equations to estimate the biomass of the more important species of the regenerating caatinga is provided. They are different from equations for mature caatinga plants of the same species.

Trees in regenerating areas have lower biomass than trees in mature areas with the same trunk diameter and height, because they have fewer, shorter or thinner branches and form smaller crowns.

**Acknowledgements**

We gratefully want to thank the following institutions that supported directly the implementation of this research: IPA headquarters and experimental stations at Sertânia and Serra Talhada; the Caroalina community for permitting us to use their land for the project; the Clothworkers Foundation and Rio Tinto plc who provided the necessary funding and CNPq with the necessary fellowships. A special thanks to all collaborators from IPA, UEFs, Kew, APNE and all the field staff that made special efforts for the success of this work.

**References**


