

Simultaneous estimation as alternative to independent modeling of tree biomass

Carlos Roberto Sanquetta¹ · Alexandre Behling¹ · Ana Paula Dalla Corte¹ · Sylvio Péllico Netto¹ · Ana Beatriz Schikowski¹ · Mauricio Koubay do Amaral¹

Received: 15 March 2014 / Accepted: 17 June 2015 / Published online: 21 July 2015
© INRA and Springer-Verlag France 2015

Abstract

• **Key message** In this paper it is shown that a simultaneous adjustment provides more efficient estimates of total tree biomass than with independent modelling for biomass estimates by compartments (canopy, bole and roots).

• **Context** When modeling tree biomass, it is important to consider the additivity property, since the total tree biomass must be equal to the sum of the biomass of the components.

• **Objective** The aim of this study was to assess the simultaneous estimation performance, considering the additivity principle with respect to independent estimate when modeling biomass components and total biomass.

Handling Editor: Erwin Dreyer

Contribution of the co-authors CRS and APDC have extensive experience in research related to biomass and have published hundreds of papers on the subject in national and international journals. They participated in the experimental design, emphasizing the importance of using a technique that addresses the problem of lack of additivity in modeling biomass.

Because of their experience in sampling theory and statistical modeling, AB and SPN conducted biomass modeling, employing both simultaneous and independent sampling methods. They also conducted a literature review on the subject.

ABS has experience in modeling biomass. She conducted the experiment and was responsible for ensuring the quality of the data. Furthermore, when independent fitting was used, lack of additivity was studied.

MA reviews of the paper.

In general, all authors were dedicated to the tasks of writing and critically revising the manuscript before the final presentation.

✉ Carlos Roberto Sanquetta
carlos_sanquetta@hotmail.com

¹ Forest Sciences Department, Forest Inventory Laboratory – Center for Excellence in Biomass Carbon Fixation Studies (BIOFIX), Federal University of Paraná, Av. Prof. Lothário Meissner, 900 - Jardim Botânico - Campus III, 80210-170 Curitiba, Paraná, Brazil

• **Methods** Individual modeling of total biomass and biomass components of leaves, branches, bole without bark, bole bark, and roots was performed on *Pinus elliottii* Engelm trees derived from forest stands in southern Brazil. Five nonlinear models were tested, and the best performance for estimating the total biomass of each component was selected, characterizing the independent estimation. The models selected for each component were fitted using the nonlinear seemingly unrelated regression method, which characterizes simultaneous estimation.

• **Results** Independent fitting of coefficients for biomass components and total biomass was not satisfactory, as the sum of the biomass component estimates diverged from the total biomass. This was not observed when the simultaneous fitting was used, which takes into account the additivity principle, and resulted in more effective estimators.

• **Conclusion** The simultaneous estimation method must be used in modeling tree biomass.

Keywords Nonlinear seemingly unrelated regressions · Additivity · Biomass components · *Pinus elliottii* · SAS

1 Introduction

Biomass is the mass of biological origin, living or dead, animal or plant, whereas the forest biomass is applicable to all existing plant mass in the forest or the arboreal fraction (Sanquetta and Balbinot 2002). The laborious and costly process of measuring tree biomass in the forest is often the motivation for using empirical models that allow biomass to be estimated from dendrometric variables. The study of the biomass modeling is associated with the subject of allometry. West et al. (1997) and Broad (1998) present a general theory for the allometric scale laws.

Scientists, in turn, are interested in knowing the partitioned biomass of the tree, namely the proportion distributed among its components: root, bole, bole bark, branches, and leaves. For this purpose, biomass measurements are performed on a per component basis and the total biomass is considered the sum of all masses. With the data in hand, the modeling study can be carried out to find the best performing equation for estimating biomass components and total biomass. Thus, the biomass of a given area can be estimated using inventory measurements of diameter at breast height and total height.

The problem proposed by the authors in this paper is to choose the best technique for building the allometric equations. When modeling the biomass components and total biomass, in most cases, the adjustment of models is performed independently. That is, isolated fitting is performed for each component. Using these equations, the sum of the biomass components does not produce the same result as that obtained using the total biomass equation. This generates inconsistent results; however, this approach was found in 80 % of 50 studies published in scientific journals in recent years.

The sum of the biomass component estimates should result in the same value as that estimated from the total biomass equation. This implies that the equations for biomass components and total biomass must be estimated together. Simultaneous estimation, considering the additivity principle of (Kozak, 1970; Reed and Green, 1985), is a technique that can be used to solve the inconsistency in these estimates.

Zellner (1962) proposed the seemingly unrelated regressions (SUR) approach, which utilizes simultaneous estimation. This technique provides a statistically correlated system of equations with restrictions (Parresol, 1999) and, in the present case, the fitting of the equations for the biomass components and for the total biomass imposes restrictions on the parameters, ensuring additivity. Specific methods to ensure additivity of regression functions are presented by Cunia (1979) and Jacobs and Cunia (1980). The SUR procedure is discussed at length by Reed (1986), Gallant (1987), and Srivastava and Giles (1987).

The plantations of *Pinus elliottii* Engelm in Brazil were established for timber production, aiming at various industrial purposes. The use of allometric equations has allowed the emergence of scientific papers related to estimates of tree biomass components (Witschoreck, 2008; Valerio, 2009; Giongo et al., 2011). However, it is not yet frequent to use equations with simultaneous adjustment for the genus *Pinus*. Recently, Dong et al. (2014) used the process nonlinear seemingly unrelated regressions (NSUR) in estimates of components of biomass for *Pinus koraiensis* Sieb. et Zucc, demonstrating success in applying this methodology.

The independent fitting accounts for the correlation between the biomass components, which ignores the principle of additivity and may result in inconsistencies in the overall estimate of the tree biomass. Thus, in this study, we tested the

following hypothesis: *Simultaneous fitting rather than independent fitting results in a more effective estimator for the equations of biomass components and total biomass.*

The objective of this study is to evaluate the performance of simultaneous estimation (which considers the principle of additivity) for the independent modeling of individual components and the total biomass of *Pinus elliottii* Engelm trees.

2 Materials and methods

2.1 Data source

In this study, 30 individuals of the species *Pinus elliottii* Engelm from forest stands in the south central region of the state of Paraná, Brazil, were used. The plantations were established in the municipality of General Carneiro, coordinates 26° 25' 44" South, 51° 19' 2" West, south central state of Paraná, Brazil. The climate of the region, according to the classification of Köppen, is Humid Subtropical Mesothermal (Cfb), characterized by cool summers and winters with frosts, no dry season. Organic soils are characteristically hydromorphic, predominantly Entisols, Inceptisols, and Ultisols. Individual trees were selected in the diameter range of 12.73 to 30.88 cm, between the ages of 6 and 20 years.

The individual biomass was measured using the destructive method. Initially the diameter at breast height (1.30 meters from ground level - cm) was measured with a dendrometric tape (Forestry) and then individuals were felled, and the total height (m) of each tree was measured using a measuring tape (Richter).

The fresh biomass was obtained from weighing the five components of each tree separately: bole without bark, bole bark, branches, leaves, and roots. Each component was weighed using a dynamometer (Instrumtemp) with a 100-kg capacity and accurate to within 100 g. To determine the root biomass, a trench was dug to the extent of the crown coverage, projected on the ground, and to a depth of 50 cm. All roots in the trench were collected greater than a diameter limit of 2 mm. The stump biomass was considered as root biomass.

A 500 g sample of each component and of each individual was collected to determine moisture content, which was obtained by measuring the fresh weight and dry weight on an analytical balance (Shimadzu). The dry weight was obtained after the samples were dried in a convection oven at 100 °C to a constant weight at 15 days. Thus, taking into consideration the moisture content of the samples, it was possible to establish a relationship to the fresh biomass, which resulted in the dry biomass of each tree's components. The total biomass was the sum of each component's masses.

A 4 cm thick disk was collected at the bole base, which was sanded to visually count growth rings and determine tree age.

This information was also compared to the historical records of the plantations.

2.2 Estimation of parameters using generalized nonlinear regression model (EGNLS)

The model used to estimate biomass components were:

$$\hat{y}_{leaves} = b_0 * d^{b_1} * \left(\frac{1}{h}\right)^{b_2} + e_i \tag{1}$$

$$\hat{y}_{branches} = b_0 * \left(\frac{1}{d}\right)^{b_1} * \left(\frac{1}{h}\right)^{b_2} + e_i \tag{2}$$

$$\hat{y}_{bole} = b_0 * (d^2 h)^{b_1} + e_i \tag{3}$$

$$\hat{y}_{bark} = b_0 * d^{b_1} + e_i \tag{4}$$

$$\hat{y}_{root} = b_0 * (d^2 h)^{b_1} + e_i \tag{5}$$

$$\hat{y}_{total} = b_0 * (d^2 h)^{b_1} + e_i \tag{6}$$

where \hat{y}_j =estimates of a given component or total biomass (in kg); d =diameter at breast height (cm); h =total height (m); b_i =coefficients of the model.

These models were selected in a previous study for the preparation of this work. The selection of the best model for the biomass of each component and total was based on the quality of fit indicators: adjusted coefficient of determination (R^2_{adj}), standard error of estimate (S_{yx} (%)) (presented by Steel et al. 1996), graphical analysis of the residuals. The White test was applied to check the homogeneity of the residuals.

The first step is to calculate the partial derivative matrix ($Z(\beta)$) defined by:

$$Z(\beta) = \frac{\partial f(X, \beta)}{\partial \beta} = \begin{bmatrix} \frac{\partial f(x_1, \beta)}{\partial \beta_1} & \dots & \frac{\partial f(x_1, \beta)}{\partial \beta_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(x_t, \beta)}{\partial \beta_1} & \dots & \frac{\partial f(x_t, \beta)}{\partial \beta_k} \end{bmatrix} \tag{7}$$

where X are the independent variables and β the coefficients of the models.

The β estimate is obtained, minimizing the residual sum of squares ($S(\beta)$):

$$S(\beta) = \varepsilon' \varepsilon = [y - f(X, \beta)]' [y - f(X, \beta)] \tag{8}$$

where y is the observed dependent variable and $f(X, \beta)$ is the estimated dependent variable.

The solution of this system is achieved by employing an algorithm; see Greene (1999) and Judge et al. (1988). This procedure was applied using proc model in *Statistical Analysis System* (SAS) software.

The confidence interval (CI) of the average value, at 95 % probability, on two side test, is given by:

$$CI \left[\hat{y} \pm t_{\alpha/2} \sqrt{\hat{v}(\hat{y})} \right] = 1 - \alpha \tag{9}$$

The variance (\hat{v}) to the mean value (\hat{y}_t) is estimated by:

$$\hat{v}(\hat{y}_t) = z(b)'_t S^2(b) z(b)_t \tag{10}$$

The variance-covariance matrix ($S^2(b)$) is given by:

$$S^2(b) = \hat{\sigma}^2 [Z(b)' Z(b)]^{-1} \tag{11}$$

The variance is given by:

$$\hat{\sigma}^2 = \frac{S(b)}{T-K} = \frac{[y - f(X, b)]' [y - f(X, b)]}{T-K} \tag{12}$$

where T is the number of observations and K is the number of parameters in the estimated model.

The prediction interval (PI) (new, i.e., n), at 95 % probability, on two side test, is calculated by:

$$PI \left[\hat{y}_{i(n)} \pm t_{\alpha/2} \sqrt{\hat{v}(\hat{y}_n)} \right] = 1 - \alpha \tag{13}$$

The variance is given by:

$$\hat{v}(\hat{y}_{i(n)}) = \hat{\sigma}^2 + z(b)'_t S^2(b) z(b)_t \tag{14}$$

2.3 Nonlinear seemingly unrelated regressions

To achieve simultaneous modeling, at first, it is necessary to select the best performing model for each component of the biomass. In the second step, the selected models for each component are adjusted simultaneously, so total biomass of the equation is a function of the independent variables in equations for each component i , thus placing restrictions on the parameters. This can be described as:

$$y_i = f_i(X_i, \beta_i) + \varepsilon_i, \quad i = 1, \dots, k \tag{15}$$

$$y_{total} = f_{total}(X_1, X_2, \dots, X_k, \beta_1, \beta_2, \dots, \beta_n) + \varepsilon_k$$

Or alternatively:

$$y_{leaves} = f_{leaves}(X, \beta) + \varepsilon_{leaves} \tag{16}$$

$$y_{branches} = f_{branches}(X, \beta) + \varepsilon_{branches}$$

$$y_{bole} = f_{bole}(X, \beta) + \varepsilon_{bole}$$

$$y_{bark} = f_{bark}(X, \beta) + \varepsilon_{bark}$$

$$y_{total} = f_{total}(X, \beta) + \varepsilon_{total}$$

The estimate of the vector β employing NSUR that corresponds to the values of β , which minimize the residual sum of squares, is given by:

$$R(\beta) = \varepsilon' \left(\sum^{-1} \otimes I \right) \varepsilon = [y-f(X, \beta)]' \left(\sum^{-1} \otimes I \right) [y-f(X, \beta)] \tag{17}$$

where $R(\beta)=r$ to impose restrictions on the parameters.

The calculation is performed using iterative convergence to minimize the residual sum of squares and can be, for example, the minimization method using the Gauss-Newton, Marquardt (SAS 1999a, b, 2002). This procedure was performed in SAS software, and the technique consisted of applying the use of generalized least squares method called *nonlinear seemingly unrelated regressions* (NSUR) iteratively using the option *ITSUR* of the procedure *MODEL* (Institute Inc 1999a, b, 2002). The application in this case is presented in Appendix.

$F(\beta)$ means the matrix of partial derivatives of the residuals on the same parameters such that:

$$F(\beta)' = \frac{\partial \varepsilon'}{\partial \beta} = \left[\frac{\partial f'_1}{\partial \beta}, \frac{\partial f'_2}{\partial \beta}, \dots, \frac{\partial f'_M}{\partial \beta} \right] \tag{18}$$

It is assumed that the vector of random variables ε follows a distribution with zero mean ($E(\varepsilon_i)=0$), constant variance and covariance, then:

$$E(\varepsilon \varepsilon') = \sum \otimes I e E(\varepsilon_i \varepsilon_j) = \sigma_{ij} I, \tag{19}$$

where \otimes is the Kronecker product.

The unknown covariances from the model established by matrix system σ_{ij} are estimated by the expression:

$$\hat{\sigma}_{ij} = \frac{1}{\sqrt{(T-K_i)} \sqrt{(T-K_j)}} e'_i e_j \tag{20}$$

where T is the number of observations. K_i and K_j are the number of parameters of each equation i and j .

Such expression provides constant variance-covariances estimates, in which $e_i=y_i-f_i(X_i, \beta_i)$ are the residuals obtained by *estimated generalized nonlinear least squares* (EGNLS)

The estimated variance-covariance matrix of the parameters is given by:

$$\hat{\sum}_b = \left[F(b)' \left(\sum^{-1} \otimes I \right) F(b) \right]^{-1} \tag{21}$$

The variance of the NSUR is given by:

$$\hat{\sigma}_{NSUR}^2 = \frac{R(b)}{MT-K} = \frac{e' \left(\sum^{-1} \otimes I \right) e}{MT-K} \tag{22}$$

where M is the number of equations used.

The estimated variance from the equation i of the system is given by:

$$S_{y_i}^2 = f_i(b)' \sum \sigma_b f_i(b) \tag{23}$$

In which $f_i(b)$ is a column vector containing the partial derivatives in order to the parameters of the equation i .

The confidence interval (CI) of a mean value, at 95 % probability, on two side test, is given by:

$$CI \left[\hat{y} \pm t_{\alpha/2} \sqrt{S_{y_i}^2} \right] = 1-\alpha \tag{24}$$

The prediction interval (new, i.e., n) (PI), at 95 % probability, on two side test, is calculated by:

$$PI \left[\hat{y}_{i(n)} \pm t_{\alpha/2} \sqrt{S_{y_i}^2 + \hat{\sigma}_{NSUR}^2 \hat{\sigma}_{ii}} \right] = 1-\alpha \tag{25}$$

2.4 Effect of independent and simultaneous adjustment in modeling the tree biomass

Comparison of the equations for the independent estimation (performed by EGNLS) and simultaneous estimation (performed by NSUR) was performed. Regarding independent estimation, the effect of lack of additivity on estimates of the overall biomass at individual and population level was evaluated. The estimate of the tree total biomass must be equal to the sum of the biomass components. This analysis was not performed for simultaneous estimation because the additivity property is already ensured.

The difference (absolute and relative) of the results generated by the total biomass equation and the sum of the results of the equations of the biomass components were verified at the individual tree level. This analysis was also performed at the stand level for ages 6, 10, 15, and 20 years. For that purpose, we have simulated the diameter at breast height and total height of the forest stands using the *SisPinus* software (Oliveira et al. 1989). The input data in the simulation were dominant height of 21 m, initial survival percentage of 95 %, and plantation with homogeneity of five. From the diameter at breast height, height, and population survival data, one can calculate the total and biomass components for an area of 1 ha.

In addition, the simulated data were used to evaluate the effect of the methods ENGLS and SUR on the estimates of the biomass components, confidence interval, and prediction interval, extrapolated to hectare.

3 Results

3.1 Estimation of parameters using nonlinear regression—estimated generalized nonlinear least squares (EGNLS)

White’s test revealed that the residuals have presented homogeneous distribution for all components (significant at 99 % probability). The equations and statistics resulting from these adjustments were:

$$\hat{y}_{leaves} = 0.070729 * d^{2.042984} * \left(\frac{1}{h}\right)^{0.4153} \quad (26)$$

$$R^2_{adj} = 0.74, S_{yx} = 26.49\%, \text{ White} = 13.97^{ns}$$

$$\hat{y}_{branches} = 0.008359 * \left(\frac{1}{d}\right)^{-2.09814} * \left(\frac{1}{h}\right)^{-0.43851} \quad (27)$$

$$R^2_{adj} = 0.71, S_{yx} = 45.00\%, \text{ White} = 17.17^{ns}$$

$$\hat{y}_{bole} = 0.005143 * (d^2 h)^{1.097758} \quad (28)$$

$$R^2_{adj} = 0.96, S_{yx} = 19.30\%, \text{ White} = 4.26^{ns}$$

$$\hat{y}_{bark} = 0.006748 * d^{2.591151} \quad (29)$$

$$R^2_{adj} = 0.92, S_{yx} = 22.96\%, \text{ White} = 6.19^{ns}$$

$$\hat{y}_{root} = 0.013898 * (d^2 h)^{0.815147} \quad (30)$$

$$R^2_{adj} = 0.90, S_{yx} = 23.15\%, \text{ White} = 6.96^{ns}$$

$$\hat{y}_{total} = 0.0289 * (d^2 h)^{0.9671} \quad (31)$$

$$R^2_{adj} = 0.96, S_{yx} = 16.77\%, \text{ White} = 4.68^{ns}$$

Incompatibility was noted between the estimations of the components and the total biomass when the equations for independent fitting were used, and consequently, one of the properties that ensure the appropriate biomass modeling was not assured. The absence of additivity represents the existence of inconsistencies in the overall estimate of the tree biomass, since the estimated values from the biomass equations for the components are not equal to that estimated from the equation for total tree biomass.

These differences inside a tree varied from -3,02 to 4.37 kg and occurred for the set of diameters at breast height and total height used when applying the equation. When the diameter at breast height and total height were simulated per hectare in pine forest stands of age 6 to 20 years, and after applying these data in the equations for total and biomass components (estimated from total biomass equations—estimated from components), the differences reached up to 7.63 % (Table 1).

A solution to force additivity of the equations solved by EGNLS, is consider the total biomass as the sum of all biomass components, namely: $\hat{y}_{total} = \hat{y}_{leaves} + \hat{y}_{branches} + \hat{y}_{bole} + \hat{y}_{bark} + \hat{y}_{root}$.

Take a tree with dbh=21 cm and ht=16 m, and by applying these values in the equations of 26–30, we have:

$$\hat{y}_{leaves} = 11.24 \text{ kg} \quad (32)$$

$$\hat{y}_{branches} = 16.76 \text{ kg} \quad (33)$$

$$\hat{y}_{bole} = 86.30 \text{ kg} \quad (34)$$

$$\hat{y}_{bark} = 18.00 \text{ kg} \quad (35)$$

$$\hat{y}_{root} = 19.06 \text{ kg} \quad (36)$$

$$\hat{y}_{total} = 11.24 + 16.76 + 86.30 + 18.00 + 19.06 = 151.36 \text{ kg} \quad (37)$$

The vector of partial derivatives resulted in:

$$z(b)'_{leaves} = \left[d^{b_1} \left(\frac{1}{h}\right)^{b_2} \quad b_0 d^{b_1} \ln(d) \left(\frac{1}{h}\right)^{b_2} \quad b_0 d^{b_1} \ln\left(\frac{1}{h}\right) \left(\frac{1}{h}\right)^{b_2} \right] \\ = [158.93 \quad 34.22 \quad -31.17] \quad (38)$$

$$z(b)'_{branches} = \left[\left(\frac{1}{d}\right)^{b_1} \left(\frac{1}{h}\right)^{b_2} \quad b_0 \left(\frac{1}{d}\right)^{b_1} \ln\left(\frac{1}{d}\right) \left(\frac{1}{h}\right)^{b_2} \quad b_0 \left(\frac{1}{d}\right)^{b_1} \ln\left(\frac{1}{h}\right) \left(\frac{1}{h}\right)^{b_2} \right] \\ = [2005.49 \quad -51.04 \quad -46.48] \quad (39)$$

$$z(b)'_{bole} = \left[(d^2 h)^{b_1} \quad b_0 (d^2 h)^{b_1} \ln(d^2 h) \right] \\ = [16,779.77 \quad 764.74] \quad (40)$$

$$z(b)'_{bark} = \left[d^{b_1} \quad b_0 d^{b_1} \ln(d) \right] = [2667.29 \quad 54.80] \quad (41)$$

$$z(b)'_{root} = \left[(d^2 h)^{b_1} \quad b_0 (d^2 h)^{b_1} \ln(d^2 h) \right] \\ = [1371.33 \quad 168.89] \quad (42)$$

Table 1 The difference (absolute and relative) of the result generated by the equation of total biomass for *Pinus elliottii* Engelm. and sum of the results of the equations for biomass components in an area of 1 ha

Age (years)	Total biomass equation (t ha ⁻¹)— 1	Σ of biomass equations per component (t ha ⁻¹)— 2	Difference (t ha ⁻¹) 1-2	Difference (%) 1-2
6	32.77	30.27	2.50	7.63
10	152.70	147.71	4.99	3.27
15	272.44	269.09	3.35	1.23
20	366.74	366.66	0.08	0.02

The variance-covariance matrix b , obtained by EGNLS regression for biomass components, resulted in:

$$S^2(b)_{\text{leaves}} = \begin{bmatrix} 0.0020404 & -0.0138537 & -0.0053848 \\ -0.0138537 & 0.1596963 & 0.1083941 \\ -0.0053848 & 0.1083941 & 0.0930543 \end{bmatrix} \quad (43)$$

$$S^2(b)_{\text{branches}} = \begin{bmatrix} 0.0001112 & 0.0039757 & 0.0000622 \\ 0.0039757 & 0.4036055 & -0.280671 \\ 0.0000622 & -0.280671 & 0.3066558 \end{bmatrix} \quad (44)$$

$$S^2(b)_{\text{bole}} = \begin{bmatrix} 0.0000111 & -0.000222 \\ -0.000222 & 0.0044433 \end{bmatrix} \quad (45)$$

$$S^2(b)_{\text{bark}} = \begin{bmatrix} 0.0000183 & -0.0008231 \\ -0.0008231 & 0.0370452 \end{bmatrix} \quad (46)$$

$$S^2(b)_{\text{root}} = \begin{bmatrix} 0.0000794 & -0.0005937 \\ -0.0005937 & 0.0044509 \end{bmatrix} \quad (47)$$

The variance for the estimation of biomass components resulted in:

$$\hat{v}(\hat{y}_{\text{leaves}}) = 0.3798 \text{ kg}^2 \quad (48)$$

$$\hat{v}(\hat{y}_{\text{branches}}) = 3.9702 \text{ kg}^2 \quad (49)$$

$$\hat{v}(\hat{y}_{\text{bole}}) = 26.4069 \text{ kg}^2 \quad (50)$$

$$\hat{v}(\hat{y}_{\text{bark}}) = 0.8215 \text{ kg}^2 \quad (51)$$

$$\hat{v}(\hat{y}_{\text{root}}) = 1.2656 \text{ kg}^2 \quad (52)$$

The total biomass is given by the sum of the biomass components ($\hat{y}_{\text{total}} = \hat{y}_{\text{leaves}} + \hat{y}_{\text{branches}} + \hat{y}_{\text{bole}} + \hat{y}_{\text{bark}} + \hat{y}_{\text{root}}$), consequently the variance is given by:

$$\hat{v}(\hat{y}_{\text{total}}) = \sum v(\hat{y}_i) + 2 \sum \sum cov(\hat{y}_i, \hat{y}_j) \quad (53)$$

In which $cov(\hat{y}_i, \hat{y}_j)$ is calculated by:

$$cov(\hat{y}_i, \hat{y}_j) = \hat{\rho}_{y_i y_j} \sqrt{v(\hat{y}_i) v(\hat{y}_j)}. \quad (54)$$

where $\hat{\rho}_{y_i y_j}$ means expected correlation between y_i and y_j , i.e.,:

Component	Leaves	Branches	Bole	Bark	Root	Total
Leaves	1					
Branches	0.8019	1				
Bole	0.7823	0.8928	1			
Bark	0.8352	0.8584	0.9666	1		
Root	0.7907	0.8162	0.9357	0.9191	1	
Total	0.8155	0.9166	0.9968	0.9725	0.9438	1

The application of the Eq. 54 resulted in:

$$\begin{aligned} \hat{v}(\hat{y}_{\text{total}}) &= 0.3798 + 3.9702 + 26.4069 + 0.8215 + 1.2656 + 2 \\ &\quad * 0.8019 \sqrt{0.3798 + 3.9702} + 2 * 0.7823 \sqrt{0.3798 + 26.4069} + 2 \\ &\quad * 0.8352 \sqrt{0.3798 + 0.8215} + 2 * 0.7907 \sqrt{0.3798 + 1.2656} + 2 \\ &\quad * 0.8928 \sqrt{3.9702 + 26.4069} + 2 * 0.8584 \sqrt{3.9702 + 0.8215} + 2 \\ &\quad * 0.8162 \sqrt{3.9702 + 1.2656} + 2 * 0.9666 \sqrt{26.4069 + 0.8215} + 2 \\ &\quad * 0.9357 \sqrt{26.4069 + 1.2656} + 2 * 0.9191 \sqrt{0.8215 + 1.2656} \\ &= 88.0684 \text{ kg}^2 \end{aligned} \quad (55)$$

Estimates of confidence intervals at 95 % probability level resulted in:

$$CI_{\text{leaves}} = 11.24 \pm 2.04 \sqrt{0.3798} = 11.24 \pm 1.26 \text{ kg} \quad (56)$$

$$CI_{\text{branches}} = 16.76 \pm 2.04 \sqrt{3.9702} = 16.76 \pm 4.06 \text{ kg} \quad (57)$$

$$CI_{\text{bole}} = 86.30 \pm 2.04 \sqrt{26.4069} = 86.30 \pm 10.48 \text{ kg} \quad (58)$$

$$CI_{\text{bark}} = 18.00 \pm 2.04 \sqrt{0.8215} = 18.00 \pm 1.85 \text{ kg} \quad (59)$$

$$CI_{\text{root}} = 19.06 \pm 2.04 \sqrt{1.2656} = 19.06 \pm 2.29 \text{ kg} \quad (60)$$

$$CI_{\text{total}} = 151,36 \pm 2.04 \sqrt{88.0684} = 151,36 \pm 19.14 \text{ kg} \quad (61)$$

Estimates of prediction intervals at 95 % probability level resulted in:

$$PI_{\text{leaves}} = 11.24 \pm 2.04 \sqrt{9.4117} = 11.24 \pm 6.26 \text{ kg} \quad (62)$$

$$PI_{\text{branches}} = 16.76 \pm 2.04 \sqrt{72.0528} = 16.76 \pm 17.32 \text{ kg} \quad (63)$$

$$PI_{\text{bole}} = 86.30 \pm 2.04 \sqrt{449.3069} = 86.30 \pm 43.24 \text{ kg} \quad (64)$$

$$PI_{\text{bark}} = 18.00 \pm 2.04 \sqrt{20.9645} = 18.00 \pm 9.34 \text{ kg} \quad (65)$$

$$PI_{\text{root}} = 19.06 \pm 2.04 \sqrt{24.7059} = 19.06 \pm 10.14 \text{ kg} \quad (66)$$

$$PI_{\text{total}} = 151.36 \pm 2.04 \sqrt{811.3462} = 151.36 \pm 58.11 \text{ kg} \quad (67)$$

3.2 Nonlinear seemingly unrelated regressions

White's test revealed that the residuals have presented homogeneous distribution for all components (significant at 99 % probability). The equations and statistics resulting from these adjustments were

$$\begin{aligned} \hat{y}_{\text{leaves}} &= 0.101855 * d^{1.523562} * \left(\frac{1}{h}\right)^{-0.02921} \\ R_{\text{adj}}^2 &= 0.73, S_{yx} = 26.91\%, \text{ White} = 15.69^{\text{ns}}. \end{aligned} \quad (68)$$

$$\hat{y}_{\text{branches}} = 0.003479 * \left(\frac{1}{d}\right)^{-2.77207} * \left(\frac{1}{h}\right)^{0.009156} \quad (69)$$

$$R_{\text{adj}}^2 = 0.72, S_{yx} = 44.45\%, \text{White} = 13.59^{\text{ns}}$$

Take the same tree with dbh=21 cm and ht=16 m, and by applying these values in the equations of 68–73, we have:

$$\hat{y}_{\text{bole}} = 0.003247 * (d^2 h)^{1.144185} \quad (70)$$

$$R_{\text{adj}}^2 = 0.96, S_{yx} = 19.16\%, \text{White} = 4.06^{\text{ns}}$$

$$\hat{y}_{\text{leaves}} = 11.42 \text{ kg} \quad (74)$$

$$\hat{y}_{\text{branches}} = 15.69 \text{ kg} \quad (75)$$

$$\hat{y}_{\text{bark}} = 0.005036 * d^{2.679999} \quad (71)$$

$$R_{\text{adj}}^2 = 0.92, S_{yx} = 22.65\%, \text{White} = 5.65^{\text{ns}}$$

$$\hat{y}_{\text{bole}} = 82.21 \text{ kg} \quad (76)$$

$$\hat{y}_{\text{raiz}} = 0.012858 * (d^2 h)^{0.82377} \quad (72)$$

$$R_{\text{adj}}^2 = 0.90, S_{yx} = 22.77\%, \text{White} = 6.88^{\text{ns}}$$

$$\hat{y}_{\text{bark}} = 17.70 \text{ kg} \quad (77)$$

$$\hat{y}_{\text{root}} = 19.03 \text{ kg} \quad (78)$$

$$\hat{y}_{\text{total}} = \hat{y}_{\text{leaves}} + \hat{y}_{\text{branches}} + \hat{y}_{\text{bole}} + \hat{y}_{\text{bark}} + \hat{y}_{\text{root}} \quad (73)$$

$$R_{\text{adj}}^2 = 0.90, S_{yx} = 19.24\%, \text{White} = 28.53^{\text{ns}}$$

$$\hat{y}_{\text{total}} = 11.42 + 15.69 + 82.21 + 17.70 + 19.03 = 145.95 \text{ kg} \quad (79)$$

The vector of the partial derivatives is given by:

$$f(b)'_{\text{leaves}} = \left[d^{b_1} \left(\frac{1}{h}\right)^{b_2} \quad b_0 d^{b_1} \ln(d) \left(\frac{1}{h}\right)^{b_2} \quad b_0 d^{b_1} \ln\left(\frac{1}{h}\right) \left(\frac{1}{h}\right)^{b_2} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \right] \quad (80)$$

$$= [112.11 \quad 34.77 \quad -31.66 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$

$$f(b)'_{\text{branches}} = \left[0 \quad 0 \quad 0 \quad \left(\frac{1}{d}\right)^{b_1} \left(\frac{1}{h}\right)^{b_2} \quad b_0 \left(\frac{1}{d}\right)^{b_1} \ln\left(\frac{1}{d}\right) \left(\frac{1}{h}\right)^{b_2} \quad b_0 \left(\frac{1}{d}\right)^{b_1} \ln\left(\frac{1}{h}\right) \left(\frac{1}{h}\right)^{b_2} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \right] \quad (81)$$

$$= [0 \quad 0 \quad 0 \quad 4510.86 \quad -47.78 \quad -43.51 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$

$$f(b)'_{\text{bole}} = \left[0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad (d^2 h)^{b_1} \quad b_0 (d^2 h)^{b_1} \ln(d^2 h) \quad 0 \quad 0 \quad 0 \quad 0 \right] \quad (82)$$

$$= [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 25,319.96 \quad 728.55 \quad 0 \quad 0 \quad 0 \quad 0]$$

$$f(b)'_{\text{bark}} = \left[0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad d^{b_1} \quad b_0 d^{b_1} \ln(d) \quad 0 \quad 0 \quad 0 \right] \quad (83)$$

$$= [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 3495.80 \quad 53.60 \quad 0 \quad 0]$$

$$f(b)'_{\text{root}} = \left[0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad (d^2 h)^{b_1} \quad b_0 (d^2 h)^{b_1} \ln(d^2 h) \right] \quad (84)$$

$$= [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1480.22 \quad 168.66]$$

The variance-covariance matrix of b_{ij} (where i refers to coefficients in each equation and j refers to coefficients among

equations) for the biomass components in the regressions, obtained by EGNLS resulted in:

$$\hat{\Sigma}_b = \begin{matrix} & \begin{matrix} b_{11} & b_{12} & b_{13} & b_{21} & b_{22} & b_{23} \end{matrix} \\ \begin{matrix} b_{11} \\ b_{12} \\ b_{13} \\ b_{21} \\ b_{22} \\ b_{23} \\ b_{31} \\ b_{32} \\ b_{41} \\ b_{42} \\ b_{51} \\ b_{52} \end{matrix} & \begin{pmatrix} 0.003751 & -0.012720 & -0.001400 & 0.000090 & 0.006978 & 0.000972 \\ -0.012720 & 0.061885 & 0.024932 & -0.000356 & -0.029000 & -0.001197 \\ -0.001400 & 0.024932 & 0.022481 & -0.000092 & -0.008714 & 0.001891 \\ 0.000090 & -0.000356 & -0.000092 & 0.000009 & 0.000797 & -0.000023 \\ 0.006978 & -0.029000 & -0.008714 & 0.000797 & 0.143586 & -0.081261 \\ 0.000972 & -0.001197 & 0.001891 & -0.000023 & -0.081261 & 0.086361 \\ -0.000011 & -0.000011 & -0.000046 & 0.000000 & 0.000027 & -0.000004 \\ 0.000356 & 0.000162 & 0.001322 & -0.000008 & -0.000780 & 0.000120 \\ 0.000011 & -0.000060 & -0.000031 & -0.000001 & -0.000088 & 0.000024 \\ -0.000622 & 0.003327 & 0.001580 & 0.000046 & 0.005442 & -0.001518 \\ -0.000049 & 0.000168 & 0.000023 & -0.000004 & -0.000196 & -0.000132 \\ 0.000399 & -0.001362 & -0.000172 & 0.000029 & 0.001542 & 0.001105 \end{pmatrix} \\ & \begin{matrix} b_{31} & b_{32} & b_{41} & b_{42} & b_{51} & b_{52} \end{matrix} \\ \begin{matrix} b_{11} \\ b_{12} \\ b_{13} \\ b_{21} \\ b_{22} \\ b_{23} \\ b_{31} \\ b_{32} \\ b_{41} \\ b_{42} \\ b_{51} \\ b_{52} \end{matrix} & \begin{pmatrix} -0.000011 & 0.000356 & 0.000011 & -0.000622 & -0.000049 & 0.000399 \\ -0.000011 & 0.000162 & -0.000060 & 0.003327 & 0.000168 & -0.001362 \\ -0.000046 & 0.001322 & -0.000031 & 0.001580 & 0.000023 & -0.000172 \\ 0.000000 & -0.000008 & -0.000001 & 0.000046 & -0.000004 & 0.000029 \\ 0.000027 & -0.000780 & -0.000088 & 0.005442 & -0.000196 & 0.001542 \\ -0.000004 & 0.000120 & 0.000024 & -0.001518 & -0.000132 & 0.001105 \\ 0.000001 & -0.000038 & 0.000000 & -0.000019 & 0.000001 & -0.000011 \\ -0.000038 & 0.001183 & -0.000010 & 0.000595 & -0.000042 & 0.000339 \\ 0.000000 & -0.000010 & 0.000009 & -0.000509 & -0.000012 & 0.000098 \\ -0.000019 & 0.000595 & -0.000509 & 0.030594 & 0.000730 & -0.005895 \\ 0.000001 & -0.000042 & -0.000012 & 0.000730 & 0.000065 & -0.000524 \\ -0.000011 & 0.000339 & 0.000098 & -0.005895 & -0.000524 & 0.004239 \end{pmatrix} \end{matrix} \tag{85}$$

The variance for the estimate of biomass components resulted in:

$$S_{y_{leaves}}^2 = f_i(b)' \hat{\Sigma}_b f_i(b) = 0.3713 \text{ kg}^2 \tag{86}$$

$$S_{y_{branches}}^2 = f_i(b)' \hat{\Sigma}_b f_i(b) = 2.1319 \text{ kg}^2 \tag{87}$$

$$S_{y_{bole}}^2 = f_i(b)' \hat{\Sigma}_b f_i(b) = 10.0356 \text{ kg}^2 \tag{88}$$

$$S_{y_{bark}}^2 = f_i(b)' \hat{\Sigma}_b f_i(b) = 1.1723 \text{ kg}^2 \tag{89}$$

$$S_{y_{root}}^2 = f_i(b)' \hat{\Sigma}_b f_i(b) = 1.4035 \text{ kg}^2 \tag{90}$$

$$S_{y_{total}}^2 = f_i(b)' \hat{\Sigma}_b f_i(b) = 15.1666 \text{ kg}^2 \tag{91}$$

Estimates of confidence intervals at 95 % probability level resulted in:

$$CI_{leaves} = 11.42 \pm 2.04\sqrt{0.3713} = 11.42 \pm 1.24 \text{ kg} \tag{92}$$

$$CI_{branches} = 15.69 \pm 2.04\sqrt{2.1319} = 15.69 \pm 2.98 \text{ kg} \tag{93}$$

$$CI_{bole} = 82.21 \pm 2.04\sqrt{10.0356} = 82.21 \pm 6.46 \text{ kg} \tag{94}$$

$$CI_{bark} = 17.70 \pm 2.04\sqrt{1.1723} = 17.70 \pm 2.21 \text{ kg} \tag{95}$$

$$CI_{root} = 19.03 \pm 2.04\sqrt{1.4035} = 19.03 \pm 2.42 \text{ kg} \tag{96}$$

$$CI_{total} = 145.95 \pm 2.04\sqrt{15.1666} = 145.95 \pm 7.94 \text{ kg} \tag{97}$$

Estimates of prediction intervals at 95 % probability level resulted in:

$$PI_{leaves} = 11.42 \pm 2.04\sqrt{1.7371} = 11.42 \pm 2.69 \text{ kg} \tag{98}$$

$$PI_{branches} = 15.69 \pm 2.04\sqrt{31.2246} = 15.69 \pm 11.40 \text{ kg} \tag{99}$$

$$PI_{bole} = 82.21 \pm 2.04\sqrt{345.1596} = 82.21 \pm 37.90 \text{ kg} \tag{100}$$

$$PI_{bark} = 17.70 \pm 2.04\sqrt{5.8521} = 17.70 \pm 4.93 \text{ kg} \tag{101}$$

$$PI_{root} = 19.03 \pm 2.04\sqrt{6.0311} = 19.03 \pm 5.01 \text{ kg} \tag{102}$$

$$PI_{total} = 145.95 \pm 2.04\sqrt{639.4817} = 145.95 \pm 51.59 \text{ kg} \tag{103}$$

3.3 Effect of independent and simultaneous adjustment in the statistics of the equations

The equations presented for estimating the biomass of individual components such as leaves, branches, bole, bark, root, and total biomass of the tree showed satisfactory levels of accuracy, regardless of the fitting method. The implementation of the simultaneous estimation method did not lead to appreciable changes on the fit’s statistics, residual distribution or regression constraints compared with those produced by independent estimation (Table 2). The values of the coefficient of determination have not presented major differences, seen only in biomass of leaf, branches, and total. The estimated standard error resulted in variations of the order of -14.73 to 1.64 %, keeping the same trends when comparing the two methods. In absolute terms, these differences were smaller than 2.47 . The distribution of residuals showed the same trend when two methods for estimation of coefficients were compared.

The largest differences were observed in the confidence intervals and prediction interval when the adjustment SUR was conducted to estimate the coefficients in a more consistent manner, since the intervals were smaller (Table 2). Similar results were observed for the estimates extrapolated to hectare values (Tables 3, 4, 5, 6, and 7). The results show that more accurate estimates are obtained by adjusting SUR, especially for the total biomass, in which the differences in confidence intervals ranged up to 71.71 % and the prediction intervals to 10.95 %. For the other biomass components, the SUR method

also provided better results, especially in the prediction intervals.

4 Discussion

The worst statistical evaluation of goodness of fit (adjusted coefficient of determination and standard error of the estimate) was observed for the biomass of leaves and branches, while the best was observed for the bole, bark, and roots. The biomass of leaves and branches is naturally more variable compared to other biomass components, since they suffer greater influences of internal and external factors, as highlighted by Satoo and Madgwick (1982) and António et al. (2007). Considering these statistics, the equations presented here for estimating the biomass of individual components such as leaves, branches, bole, bark of the bole and roots, and also for the total biomass, yielded satisfactory accuracy level. Simultaneous fitting was found not to lead to appreciable changes in the fit statistics of the equations when compared to independent fitting. It was found that the simultaneous adjustment (NSUR) did not lead to appreciable variations in the statistics of adjustments by equations, compared with independent fitting (EGNLS), with exception of the confidence intervals and prediction intervals, where the NSUR adjustment was more efficient.

Modeling biomass components and total biomass with an isolated equation for each one, as happens when using the EGNLS procedure, should not be the main concern in studies of modeling of biomass. The equation for estimation of

Table 2 Effect of independent (ENGLS) and simultaneous (NSUR) fitting on statistical estimates using models for biomass components and total biomass of individuals of *Pinus elliottii* Engelm.: leaves, branches, bole without bark, bole bark, root and total

Component	R^2_{adj}				Syx %			
	ENGLS	NSUR	AD	PD (%)	ENGLS	NSUR	AD	PD (%)
Leaves	0.74	0.73	0.01	1.35	26.49	26.91	-0.42	-1.59
Branches	0.71	0.72	-0.01	-1.41	45.00	44.45	0.55	1.22
Bole	0.96	0.96	0.00	0.00	19.30	19.16	0.14	0.73
Bark	0.92	0.92	0.00	0.00	22.96	22.65	0.31	1.35
Root	0.90	0.90	0.00	0.00	23.15	22.77	0.38	1.64
Total	0.96	0.90	0.06	6.25	16.77	19.24	-2.47	-14.73
Component	CI				PI			
	ENGLS	NSUR	AD	PD (%)	ENGLS	NSUR	AD	PD (%)
Leaves	1.26	1.24	0.02	1.50	6.26	2.69	3.57	57.02
Branches	4.06	2.98	1.08	26.60	17.32	11.40	5.92	34.18
Bole	10.48	6.46	4.02	38.36	43.24	37.90	5.34	12.35
Bark	1.85	2.21	-0.36	-19.46	9.34	4.93	4.41	47.22
Root	2.29	2.42	-0.13	-5.68	10.14	5.01	5.13	50.59
Total	19.14	7.94	11.20	58.52	58.11	51.59	6.52	11.22

Differences were calculated for the independent fitting, for example: $CI AD_{Total} = 19.87 - 7.94 = 11.93$ and $CI PD_{Total} = \frac{11.93}{19.87} \cdot 100 = 60.04\%$

R^2_{adj} adjusted coefficient of determination, S_{yx} standard error of the estimate, CI confidence intervals, PI prediction intervals, AD absolute difference, PD percent difference

Table 3 Simulation the effect of independent (ENGLS) and simultaneous (NSUR) fitting on confidence intervals and prediction intervals using models for biomass components and total biomass of individuals of *Pinus elliottii* Engelm.: leaves, branches, bole without bark, bole bark, root and total

Variables	Component	Average tree biomass (kg)			Biomass per hectare (t)		
		Biomass	CI	PI	Biomass	CI	PI
ENGLS							
Age=6 years N/ha=1.055 d=12.30 cm h=8.1 m	Leaves	5.00	1.37	6.28	5.27	1.45	6.63
	Branches	4.05	2.74	17.05	4.27	2.89	17.99
	Bole	12.63	4.49	42.19	13.33	4.74	44.51
	Bark	4.50	1.39	9.26	4.75	1.47	9.77
	Root	4.57	1.58	10.00	4.83	1.67	10.55
	Total	30.75	12.55	57.05	32.45	13.24	60.19
NSUR							
	Leaves	4.96	1.38	2.753	5.23	1.45	2.90
	Branches	3.58	1.65	11.126	3.78	1.74	11.74
	Bole	11.09	2.14	37.406	11.70	2.26	39.46
	Bark	4.20	1.24	4.583	4.43	1.31	4.84
	Root	4.50	1.53	4.648	4.75	1.62	4.90
	Total	28.33	3.55	51.096	29.89	3.75	53.91

CI confidence intervals, PI prediction intervals

each biomass component refers to a single mathematical relationship and that proved to be inappropriate, since the estimates obtained from equations for biomass components are not equal to the estimated ones from the equation for the total tree biomass, that is, they are not additive. In the present work, the incompatibility of the equation results for the components and for the total biomass was observed when using independent adjustment and also the spread of the error, when expanded to the area of one hectare (item 3.1).

One solution to force additivity in these equations would be to consider the total biomass as sum of the biomass of all components. However, this is not appropriate because the estimation in conditional models is the average value of the dependent variable as a function of the independent variables, while the value of any dependent variable should consider all the variables involved.

The NSUR procedure was effective in achieving an additive system of equations. This procedure considers the contemporaneous correlations between the biomass components,

Table 4 Simulation the effect of independent (ENGLS) and simultaneous (NSUR) fitting on confidence intervals and prediction intervals using models for biomass components and total biomass of individuals of *Pinus elliottii* Engelm.: leaves, branches, bole without bark, bole bark, root and total

Variables	Component	Average tree biomass (kg)			Biomass per hectare (t)		
		Biomass	CI	PI	Biomass	CI	PI
ENGLS							
Age=10 years N/ha=1.049 d=21.10 cm h=14.9 m	Leaves	11.69	1.28	6.26	12.26	1.34	6.57
	Branches	16.41	4.41	17.40	17.22	4.62	18.25
	Bole	80.64	10.44	43.23	84.60	10.96	45.35
	Bark	18.22	1.84	9.34	19.11	1.93	9.80
	Root	18.12	2.31	10.14	19.01	2.42	10.64
	Total	145.09	19.29	58.13	152.20	20.23	60.98
NSUR							
	Leaves	15.44	1.42	2.78	16.19	1.49	2.91
	Branches	26.83	3.48	11.54	28.15	3.65	12.11
	Bole	159.83	8.05	38.20	167.66	8.45	40.07
	Bark	29.62	2.63	5.14	31.07	2.75	5.39
	Root	30.72	2.54	5.07	32.22	2.67	5.32
	Total	262.44	9.03	51.77	275.29	9.47	54.30

CI confidence intervals, PI prediction intervals

Table 5 Simulation the effect of independent (ENGLS) and simultaneous (NSUR) fitting on confidence intervals and prediction intervals using models for biomass components and total biomass of individuals of *Pinus elliottii* Engelm.: leaves, branches, bole without bark, bole bark, root and total

Variables	Component	Average tree biomass (kg)			Biomass per hectare (t)		
		Biomass	CI	PI	Biomass	CI	PI
Age=15 years N/ha=1.020 d=25.50 cm h=19.4 m	ENGLS						
	Leaves	15.43	1.40	6.29	15.74	1.43	6.42
	Branches	27.41	4.15	17.34	27.96	4.23	17.68
	Bole	163.30	8.43	42.79	166.57	8.60	43.65
	Bark	29.77	1.17	9.23	30.36	1.19	9.41
	Root	30.60	2.14	10.11	31.22	2.19	10.31
	Total	266.52	17.12	57.70	271.85	17.47	58.85
	NSUR						
	Leaves	15.44	1.42	2.78	15.75	1.45	2.83
	Branches	26.83	3.48	11.54	27.37	3.55	11.77
	Bole	159.83	8.05	38.20	163.02	8.21	38.97
	Bark	29.62	2.63	5.14	30.21	2.68	5.24
	Root	30.72	2.54	5.07	31.33	2.60	5.17
	Total	262.44	9.03	51.77	267.68	9.21	52.80

CI confidence intervals, PI prediction intervals

considering the interdependence of the components so as to admit a set of allometric relations and also impose restrictions to the coefficients, which ensures the additivity of the equations. Thus, the calculation of the total biomass is provided by three factors: use of the same independent variables for the estimate of each biomass component, the addition of the best regression functions of each component and finally force the regressions for individual components estimate the total biomass (Cunia and Briggs 1985; Parresol, 1999).

This method ensures consistency between estimation of total biomass and the estimates of biomass components, thus is the appropriate way to model the total biomass and of the biomass components of the tree. This approach has been discussed for about half a century (Kozak, 1970; Chiyenda and Kozak, 1984; Cunia and Briggs, 1984, 1985, Reed and Green, 1985), although often it has been ignored. Few researchers who adopt the technique, such as: Parresol (1999, 2001), Bi et al. (2001), Návar et al. (2002), Carvalho and

Table 6 Simulation the effect of independent (ENGLS) and simultaneous (NSUR) fitting on confidence intervals and prediction intervals using models for biomass components and total biomass of individuals of *Pinus elliottii* Engelm.: leaves, branches, bole without bark, bole bark, root and total

Variables	Component	Average tree biomass (kg)			Biomass per hectare (t)		
		Biomass	CI	PI	Biomass	CI	PI
Age=20 years N/ha=972 d=28.20 cm h=22.6 m	ENGLS						
	Leaves	17.79	1.90	6.42	17.29	1.85	6.24
	Branches	36.21	5.62	17.75	35.19	5.46	17.25
	Bole	240.85	9.08	42.92	234.11	8.82	41.72
	Bark	38.64	1.45	9.27	37.55	1.41	9.01
	Root	40.84	2.94	10.31	39.70	2.86	10.02
	Total	374.32	19.03	58.04	363.83	18.50	56.42
	NSUR						
	Leaves	18.08	1.90	3.05	17.57	1.85	2.96
	Branches	35.42	4.51	11.89	34.43	4.38	11.56
	Bole	239.63	10.58	38.81	232.92	10.28	37.73
	Bark	38.79	3.37	5.55	37.71	3.28	5.40
	Root	41.11	3.45	5.58	39.96	3.36	5.43
	Total	373.03	11.31	52.21	362.59	10.99	50.75

CI confidence intervals, PI prediction intervals

Table 7 Comparison the effect of independent (ENGLS) and simultaneous (NSUR) fitting on confidence intervals and prediction intervals using models for biomass components and total biomass of individuals of *Pinus elliottii* Engelm.: leaves, branches, bole without bark, bole bark, root and total

Idade (anos)	Component	Average tree biomass (kg)		Biomass per hectare (t)		PD	
		AD		AD			
		IC	IP	IC	IP	IC (%)	IP (%)
6	Leaves	0.00	3.53	0.00	3.72	-0.18	56.18
	Branches	1.09	5.93	1.15	6.25	39.92	34.76
	Bole	2.35	4.79	2.48	5.05	52.37	11.34
	Bark	0.15	4.68	0.16	4.93	10.96	50.51
	Root	0.05	5.35	0.05	5.65	2.87	53.52
	Total	9.00	5.95	9.50	6.28	71.71	10.44
10	Leaves	-0.15	3.49	-0.15	3.66	-11.56	55.66
	Branches	0.92	5.86	0.97	6.15	20.96	33.67
	Bole	2.39	5.03	2.51	5.28	22.92	11.63
	Bark	-0.78	4.20	-0.82	4.41	-42.61	45.01
	Root	-0.23	5.07	-0.25	5.32	-10.11	49.99
	Total	10.26	6.37	10.76	6.68	53.17	10.95
15	Leaves	-0.02	3.51	-0.02	3.58	-1.40	55.85
	Branches	0.67	5.79	0.68	5.91	16.04	33.43
	Bole	0.38	4.59	0.39	4.68	4.52	10.72
	Bark	-1.46	4.09	-1.49	4.18	-124.78	44.36
	Root	-0.40	5.03	-0.41	5.13	-18.69	49.81
	Total	8.09	5.93	8.25	6.05	47.26	10.28
20	Leaves	0.00	3.37	0.00	3.28	0.05	52.52
	Branches	1.11	5.86	1.08	5.69	19.80	32.99
	Bole	-1.50	4.11	-1.46	3.99	-16.52	9.57
	Bark	-1.92	3.72	-1.87	3.61	-132.66	40.09
	Root	-0.51	4.72	-0.50	4.59	-17.34	45.81
	Total	7.72	5.83	7.50	5.67	40.57	10.05

Differences were calculated for the independent fitting, for example: $CI AD_{Total\ 6\ years} = 12.55 - 3.55 = 9.00$ and $CI PD_{Total\ 6\ years} = \frac{9.00}{12.55} 100 = 71.71\%$
CI confidence intervals, *PI* prediction intervals, *AD* absolute difference, *PD* percent difference

Parresol (2003), Bi et al. (2004), Cháidez et al. (2004), Nívar et al. (2004), Lambert et al. (2005), Nívar (2009) Bi et al. (2010), Goicoa et al. (2011), Ruiz-Peinaldo et al. (2012) and Menéndez-Miguélez et al. (2013).

Despite the mathematical complications of NSUR estimates, it can be easily done using the SAS software and brings the following advantages: more consistent biomass estimates for the total and components, more consistent coefficients, and with smaller confidence and prediction limits (Cunia and Briggs, 1984; Parresol, 1999; Carvalho and Parresol, 2003). The positive effects of decreases on confidence intervals and of prediction are associated with smaller variances obtained by applying the method of simultaneous estimation, considering the contemporaneous correlation between the components. Thus, the smaller the confidence intervals for the equations, the better are the intervals of the predictions, thus generating more efficient estimators. This results in important

implications in the forest, especially in terms of biomass inventory and preparation of management plans.

In this work, it was found that the simultaneous estimation (NSUR) presents the best properties for building allometric equations for estimation of total and biomass of components, when compared with the independent estimation (EGNLS). These properties are related to the compatibility of estimated values for the biomass components and for the total (the additivity property) and generate more efficient estimators, as was discussed earlier. Consequently, the hypothesis tested in the present study was confirmed.

This study is useful for understanding the basic principles of this method and also on the used analytical procedures. Meanwhile, for a more comprehensive explanation of the statistical theory involved, one should consult Zellner (1962), Kozak (1970), Gallant (1975), Cunia (1979), Jacobs and Cunia (1980), Chiyenda (1983), and Chiyenda and Kozak

(1984), Cunia and Briggs (1984, 1985), Amemiya (1985), Reed and Green (1985), Srivastava and Giles (1987), Judge et al. (1988), Greene (1999), Parresol (1999, 2001).

5 Conclusions

Modeling total biomass and components should not be treated separately, as is the case of independent estimation. Such approach has been shown to be inappropriate.

Rather, the appropriate approach to model biomass components and total biomass of the trees should take into account the interdependence of the components by including a series of allometric relationships. This is the case when using simultaneous fitting.

In modeling biomass components and total biomass, the additivity of equations must be considered, so that the equation for total biomass is a function of the sum of the regression components.

Simultaneous estimation in modeling the total biomass and of the biomass components brings several advantages, notably the compatibility between the estimates of total biomass and components. Additionally, the coefficients can gain greater consistency.

Simultaneous fitting was found not to lead to appreciable changes in the fit statistics of the equations when compared to independent fitting, yet another reason to use this method.

Simultaneous estimation presents properties best suited for building allometric equations to estimate the biomass of the components and total biomass over independent estimation. It is therefore the method that should be used.

Appendix

Dendrometric variables of individuals of the species *Pinus elliottii* Engelm. from stands established in the southern center part of Parana state, Brazil and NSUR procedure for estimating biomass equations.

Data ex;

Input Age d ht Bole Leaves Branches Root Bark Total; ddht=(d**2)*ht; Cards;

6	13.37	6.76	14.09	5.01	3.89	9.82	4.26	37.07
6	12.73	7.55	12.55	6.48	7.13	4.35	1.94	32.45
6	13.37	7.52	14.04	9.51	5.23	7.06	3.93	39.78
6	14.32	9.56	17.14	6.46	7.14	6.06	2.79	39.58
6	14.32	7.30	15.49	4.89	9.24	7.32	3.24	40.18
6	15.60	7.69	19.60	7.57	5.44	12.12	4.13	48.86
7	15.60	9.39	28.74	12.27	8.84	14.38	6.35	70.57
8	17.83	10.75	36.26	5.65	4.90	8.17	9.59	64.58

8	17.03	10.62	28.65	8.04	4.24	10.44	9.31	60.67
9	19.10	14.86	59.85	10.05	19.10	16.20	15.07	120.26
9	14.16	13.30	31.69	5.69	6.36	8.48	14.09	66.32
9	16.23	11.00	35.70	9.55	11.04	10.52	6.35	73.15
9	17.98	16.58	55.88	6.30	7.28	13.21	9.54	92.21
10	17.83	12.63	40.45	9.43	6.12	11.44	10.62	78.06
10	20.69	13.60	60.96	11.07	8.30	20.20	11.32	111.85
12	21.49	15.70	83.49	13.34	25.14	13.89	18.79	154.66
12	23.24	18.60	113.17	11.31	9.37	27.92	19.35	181.13
12	21.96	15.57	82.24	12.63	14.27	21.62	18.75	149.51
14	28.97	20.81	240.68	17.24	38.81	42.36	42.20	381.28
16	20.05	22.00	183.28	8.42	42.07	19.90	23.91	296.28
16	26.10	22.30	175.05	20.81	30.77	27.72	39.96	294.32
16	28.97	22.78	263.14	18.22	46.12	44.11	36.94	408.52
16	29.60	19.10	212.94	18.17	42.80	52.84	39.15	365.90
19	22.28	19.93	124.71	16.92	21.79	23.29	28.69	215.41
19	16.87	16.90	56.99	4.12	10.34	6.16	13.94	91.56
19	17.51	19.12	66.33	3.99	6.63	11.03	12.92	100.91
19	28.97	24.10	257.12	19.52	23.62	38.82	47.30	386.38
19	29.60	23.76	257.83	13.70	25.57	50.22	40.30	387.63
20	28.65	24.49	335.56	15.56	42.73	47.19	47.29	488.32
20	30.88	22.10	273.30	28.42	55.77	40.48	44.52	442.49

```
Proc Model Method=Mquardt;
Leaves=b11*(d**b12)*(1/(ht)**b13);
Branches=b21*(1/d)**b22;
Bole=b31*(ddht**b32);
Bark=b41*d**b42;
Root=b51*(ddht**b52);
Total = ( b 1 1 * ( d * * b 1 2 ) * ( 1 / ( h t ) * * b 1 3 ) ) +
(b21*(1/d)**b22)+(b31*(ddht**b32))+(b41*d**b42)+
(b51*(ddht**b52));
Fit Leaves Branches Bole Bark Root Total / itsur; Run
```

References

Amemiya T (1985) Advanced econometrics. Harvard University Press, Cambridge

Antônio N, Tomé M, Tomé J, Soares P, Fontes L (2007) Effect of tree, stand, and site variables on the allometry of *Eucalyptus globulus* tree biomass. *Can J For Res* 37:895–906. doi:10.1139/X06-276

Bi H, Birk E, Turner J, Lambert M, Jurskis V (2001) Converting stem volume to biomass with additivity, bias correction, and confidence bands for two Australian tree species. *N Z J For Sci* 31:298–319

Bi H, Turner J, Lambert MJ (2004) Additive biomass equations for native eucalypt forest trees in temperate Australia. *Trees* 18:467–479. doi:10.1007/s00468-004-0333-Z

Bi H, Long Y, Turner J, Lei Y, Snowdon P, Li Y, Harper R, Zerihun A, Ximenes F (2010) Additive prediction of aboveground biomass for *Pinus radiata* (D. Don) plantations. *Forest Ecol Manag* 259:2301–2314. doi:10.1016/j.foreco.2010.03.003

Broad LR (1998) Allometry and growth. *For Sci* 44:458–464

Carvalho JP, Parresol BR (2003) Additivity in tree biomass components of Pyrenean oak (*Quercus pyrenaica* Willd.). *Forest Ecol Manag* 179:269–276. doi:10.1016/S0378-1127(02)00549-2

- Cháidez JN, Barrientos NG, Luna JG, Dale V, Parresol BR (2004) Additive biomass equations for pin species of forest plantations of Durango, Mexico. *Madera y Bosques* 10:17–28
- Chiyenda SS (1983) Additivity of component regression equations when the underlying model is linear. Thesis, University of British Columbia
- Chiyenda SS, Kozak A (1984) Additivity of component biomass regression equations when the underlying model is linear. *Can J For Res* 14:441–446
- Cunia T (1979) On tree biomass tables and regression: some statistical comments. In: Frayern WE (ed) Forest resource inventory workshop proceedings. Colorado State University, Fort Collins
- Cunia T, Briggs RD (1984) Forcing additivity of biomass tables—some empirical results. *Can J For Res* 14:376–384
- Cunia T, Briggs RD (1985) Forcing additivity of biomass tables: use of the generalized least squares method. *Can J For Res* 15:23–28
- Dong L, Zhang L, Li F (2014) A compatible system of biomass equations for three conifer species in Northeast, China. *Forest Ecol Manag* 329:306–317. doi:10.1016/j.foreco.2014.05.050
- Gallant AR (1975) Seemingly unrelated non-linear regressions. *J Econometrics* 3:35–50
- Gallant AR (1987) Nonlinear statistical models. Wiley, New York
- Giongo M, da Silva DB, Koehler HS, Marchetti M (2011) Inventário de biomassa em um plantio de *Pinus elliottii* engelm. aos 23 anos de idade. *J Biotec Biodivers* 2:81–86
- Goicoa T, Militino AF, Ugarte MD (2011) Modelling aboveground tree biomass while achieving the additivity property. *Environ Ecol Estat* 18:367–384. doi:10.1007/s10651-0100137-9
- Greene WH (1999) Econometric analysis. Prentice-Hall, Upper Saddle River, 1004 pp
- Jacobs MW, Cunia T (1980) Use of dummy variables to harmonize tree biomass tables. *Can J For Res* 10:483–490
- Judge GG, Hill RC, Griffiths WE, Lütkepohl H, Lee TC (1988) Introduction to the theory and practice of econometrics. Wiley, New York
- Kozak A (1970) Methods of ensuring additivity of biomass components by regression analysis. *For Chron* 46:402–404
- Lambert MC, Ung CH, Raulier F (2005) Canadian national tree above-ground biomass equations. *Can J For Res* 35:1996–2018. doi:10.1139/X05-112
- Menéndez-Miguélez M, Canga E, Barrio-Anta M, Majada J, Álvarez-Álvarez P (2013) A three level system for estimating the biomass of *Castanea sativa* Mill. coppice stands in north-west Sapin. *Forest Ecol Manag* 291:417–426. doi:10.1016/j.foreco.2012.11.040
- Návar J (2009) Biomass component equations for Latin American species and groups species. *Ann For Sci* 66:1–8. doi:10.1051/forest/2009001
- Návar J, Méndez E, Dale V (2002) Estimating stand biomass in the Tamaulipan thornscrub of northeastern Mexico. *Ann For Sci* 59: 813–821. doi:10.1051/forest:2002079
- Návar J, Méndez E, Nájera A, Graciano J, Dale V, Parresol B (2004) Biomass equations for shrub species of Tamaulipan thornscrub of North-eastern Mexico. *J Arid Environ* 59:657–674. doi:10.1016/j.jaridenv.2004.02.010
- Oliveira YMM, Oliveira EB, Hafley WL (1989) SisPinus: simulador de crescimento e de produção para plantios de *Pinus elliottii* e *Pinus taeda* sob manejo no sul do Brasil. In: Encontro Brasileiro de planejamento florestal, 1. 1989. EMBRAPA-CNPQ, Curitiba, pp 107–118
- Parresol BR (1999) Assessing tree and stand biomass: a review with examples and critical comparisons. *For Sci* 45:573–593
- Parresol BR (2001) Additivity of nonlinear biomass equations. *Can J For Res* 31:865–878. doi:10.1139/cjfr-31-5-865
- Reed DD (1986) Estimation procedures for analytically related growth and yield models. In: Proceedings of the 1986 Annual Meeting of the Society of American Foresters, pp. 89–92
- Reed D, Green EJ (1985) A method of forcing additivity of biomass tables when using nonlinear models. *Can J For Res* 15:1184–1187
- Ruiz-Peinaldo R, Montero G, Del Rio M (2012) Biomass models to estimate carbon stocks for hardwood tree species. *For Systems* 21: 42–52. doi:10.5424/fs/2112211-02193
- Sanquetta CR, Balbinot R (2002) Metodologias para determinação de biomassa florestal. In: As Florestas e o Carbono, pp. 77–92
- SAS Institute Inc (1999a) SAS/STAT User's Guide, Version 8.2. SAS Institute, Cary Inc
- SAS Institute Inc (1999b) SAS/STAT User's Guide, Version 8. SAS Institute, Cary Inc
- SAS Institute Inc (2002) SAS/STAT User's Guide, Version 9.2. SAS Institute, Cary Inc
- Satoo T, Madgwick HA (1982) Forest biomass. Martinus Nijhoff, The Hague
- Srivastava VK, Giles DEA (1987) Seemingly unrelated regression equations models. Estimation and inference. Marcel Decker, New York
- Steel RGD, Dickey DA, Torrie JH (1996) Principles and procedures of statistics: a biometrical approach, 3 rdth edn. McGraw-Hill, New York
- Valerio AF (2009) Qualificação e modelagem da biomassa e carbono em plantações de *Pinus elliottii* Eng. com diferentes idades. Dissertação. Universidade Estadual do Centro-Oeste
- West GB, Brown JH, Enquist BJ (1997) A general model for the origin of allometric scaling laws in biology. *Science* 276:122–126. doi:10.1126/science.276.5309.122
- Witschoreck R (2008) Biomassa e nutrientes no corte raso de um povoamento de *Pinus taeda* L. de 17 anos de idade no município de Cambará do Sul – RS. Dissertação. Universidade de Santa Maria
- Zellner A (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J Am Stat Assoc* 57:348–368