academicJournals

Vol. 12(6), pp. 382-388, 9 February, 2017
DOI: 10.5897/AJAR2016.11897
Article Number: 398099662671
ISSN 1991-637X

http://www.academicjournals.org/AJAR

African Journal of Agricultural Research

Full Length Research Paper

Evaluation of nonlinear econometric models to estimate the wood volume of amazon forests

Antônio Cordeiro de Santana[1]*, Ádamo Lima de Santana[2], Mário Miguel Amin[3] and Nilson Luiz Costa[4]

[1]Universidade Federal Rural da Amazônia, UFRA, Brasil.
[2]Universidade Federal do Pará, UFPA, Brasil.
[3]Universidade da Amazônia, UNAMA, Brasil.
[4]Universidade Federal de Santa Maria, UFSM, Brasil

Received 1 November, 2016; Accepted 18 January, 2017

Effective timber extraction from primary managed forests requires accurate estimates of the annual volumes of wood. Unlike other studies, this paper uses dummy variables to capture the effects of atypical observations and to estimate the volumetric equations using nonlinear least squares. The results show that the artificial variables are important for the model's specification, and that the use of nonlinear least squares estimates provides more robust results for timber volumes than those generated by the ordinary least squares method. This methodology can be adopted to estimate the volume of wood in the primary Amazon forests that are selected for forest concessions.

Key words: Forest management, volume equation, timber company.

INTRODUCTION

The timber industry in the state of Pará, Brazil is currently going through an important restructuring process, in which the main requirement is managed wood extraction from primary forests for industrial processing (Santana et al., 2016). To maintain their competitiveness, timber companies need to apply rational technologies for the extraction and processing of wood; hence, there is a need to improve the techniques used in estimating the volume of timber to be harvested in each annual production unit (APU), as part of a sustainable forest management plan determined by the companies.

In this context, as we aim to generate data with high statistical significance (Silva et al., 2011; Santos et al., 2014), quantitative methods are used to estimate the volume of timber to be removed from each APU, which depends directly on the forest inventory applied in areas defined by representative samples.

The models used to estimate the volume of wood from primary forests follows those traditionally applied to data from planted forests. Along these lines, several studies conclude that the logarithmic model of the timber volume (dependent variable), and diameter of logs and tree height (independent variables), also known as the Schumacher-Hall (1933) model, has proven to be most

*Corresponding author. E-mail: acsufra@gmail.com.

adequate (Silva 1984; Milk and Regazzi, 1992; Rolim et al., 2006; Thaines et al., 2010; Silva et al., 2011). In primary forests with a high diversity of species, especially in the Amazon, many atypical values (outliers) can be observed; this considerably impairs the samples as a whole and their impact cannot be simply removed from the dataset (Santos et al., 2014). These atypical values, according to Silva et al. (2011), can introduce problems of heteroscedasticity in the residuals of the equation that, as Gujarati (2003) indicates, make parameter estimates of the volume equation inefficient and the model inadequate for making accurate inferences regarding the estimated volume.

Several alternative techniques can be adopted to mitigate this statistical problem. One of these techniques is the correction of heteroscedasticity via White's generalized method, as in Silva et al. (2011), the only known study that has applied this technique. Another way to avoid heteroscedasticity problems is to conduct logarithmic transformation of the equation (Gujarati, 2003) in order to reduce the difference in the measurement scale of the variables. This technique, however, cannot eliminate the problem when heteroscedasticity is produced by outliers. Moreover, it should be noted that the homoscedasticity hypothesis is commonly violated in samples of cross-sectional data (*cross-section*), and due to the problems caused to the estimators generated by the ordinary least squares (OLS) method, should be tested and, if necessary, corrections must be made prior to the model estimation.

The simple removal of trees that generate atypical values can introduce errors in specifying models. As a result, the residuals may not have a normal distribution and the statistics (R^2 and the standard error of the regression) used to measure the performance of the adjustment of equations may present bias. In this case, an efficient technique to solve the problem is the use of dummy variables (DVs) to isolate the effects of the error variance generated by the observations with values above or below the average of the samples.

The differential of this method in relation to those used previously is that the DV captures the magnitude of the effects that such information causes in the volume estimates, thereby correcting the errors in the model specification. Additionally, before considering the DV as relevant in the model, a test should be conducted to assess its contribution (Gujarati, 2003).

Finally, estimation using the Schumacher-Hall (1933) model is conducted via OLS, after the original equation has been linearized within the parameters via log application; the model is conventionally specified in logarithmic form according to the data sample characteristics in the form of an inverted "J". This model, however, entails inconvenience with respect to the requirements for calculating the antilogarithm of the variables to estimate its volume and respective inference. Thus, application of a correction factor to the data is required, as indicated by Leite and Regazzi (1992). In this model, due to the application of a logarithm to the dependent variable, the R^2 cannot be directly compared to models in which this convention is not applied (Gujarati, 2003). Nevertheless, we find that these issues are often neglected in volumetric studies, thereby contributing to the production of spurious and unfavourable economic results for lumbermen.

As a novel alternative to parameter estimation of the original equation without a logarithmic transformation, this paper proposes applying the nonlinear least squares (NLS) method to the original functional form of the equation, which is not linear in the parameters, in order to prevent the direct application of OLS and thus minimize the sum of squared residuals.

In this context, the objective of the study is to specify atypical values using the DV to estimate the equation using NLS and OLS, and to compare the results between models with and without DV. Furthermore, the paper verifies the performance between the estimated OLS and NLS models in generating more precise estimates of timber volumes for the forest species in the inventoried area.

MATERIALS AND METHODS

The data used in estimating the volume equations were obtained from trees harvested in a timber logging area of 95.91 ha of the APU (APU No. 11, work unit—UT No. 4, located in a dense forest Ombrófila) of the River Capim, in the Cikel in Brazil, with headquarters in the municipality of Paragominas, State of Pará, Brazil. The climate and soil characteristics of the study area can be observed in Silva et al. (2011).

The species were selected based on their economic importance to the company Cikel. The data collected for each sample tree harvested were the actual volume, based on the measurement of the circumference of the sections for every 2 m; diameter at breast height (DBH), calculated from equation 01; and commercial real bole height.

$$DBH = \frac{CAP}{\pi} \; (\pi = 3{,}141592) \tag{1}$$

In total, 234 trees were used as a sample, and the real volume (in m^3) was determined using equation 2 (Silva et al., 1984):

$$Vol = \sum_{i=1}^{k} \frac{g_i + g_{i+1}}{2} l_i \tag{2}$$

Where: *Vol* = total volume; g_i is basal area in the *i*-th position; and l_i is the length of the section in the *i*-th position. The trees' distribution by diameter class can be found in Smith et al. (2011).

Nonlinear regression model

The estimation of linear regression and nonlinear models applied to cross-sectional data is subject to the normality hypothesis of the error term, homoscedasticity and multicollinearity, being met. The first hypothesis is directly dependent on the R^2 and standard error regression statistics. The second, when violated, affects the efficiency property of the OLS estimators and makes the inference of data analysis spurious. The third, in turn, compromises the isolated contribution of each independent variable to the dependent variable, in addition to providing unreliable estimates of the volume intervals.

In this study, the normality hypothesis is evaluated using the Jarque and Bera (1987) test, the homoscedasticity hypothesis is tested via the White (1980) generalized test, and the multicollinearity hypothesis is checked using the variance inflation factor. Further description of the tests applied can be found in Gujarati (2003), Stata12 (2011) and Schwert (2009).

Following the incorporation of atypical values through DVs, the Chow (1960) test was applied to assess their relevance in the model specification, and the *t* statistic was used to confirm the importance of its effects on the OLS and NLS estimators. The formula is as follows:

$$F_c = \frac{(SQR_r - SQR_i)/m}{SQR_i/(n-k)} \quad (3)$$

Where F_c is the adapted Chow test; SQR_r and SQR_i are the sum of squared residuals of the restricted (without dummy) and unrestricted (with dummy) regressions, respectively; *m* is the number of DVs; *n* is the number observations; and *k* is the number of estimated parameters.

Finally, comparison between the estimated models is given by the standard error of regression, applying the Furnival Index (FI) (Furnival, 1961), as follows:

$$FI = antln\left(\frac{\sum lnVol_i}{n}\right).EP_r.\exp\left(\frac{n-k}{2n}\right) \quad (4)$$

Where EP_r is the standard error of the regression.

The applying a nonlinear generalized estimation method of volumetric equations requires special treatment. For this purpose, a description of the general process of estimating a nonlinear regression using NLS is presented. Thus, Equation 5 is as follows:

$$y_i = f(x_i, \beta) + \varepsilon_i, \text{ whit } \varepsilon_i \sim iidN(0, \sigma^2), (i = 1, ..., n) \quad (5)$$

Where *f* is the overall function of the x_i variables and β parameters. The method of OLS estimates the value of the parameters that minimize the sum of the squared residuals, given that the model is specified in a linear form within its parameters. In this case, the derivative of *f* with respect to the parameters does not depend of β. Therefore, the process of estimating the β parameters is intended to minimize the sum of squared residuals, given by Equation 6.

$$SQR(\beta) = \sum_i [y_i - f(x_i, \beta)]^2 \quad (6)$$

The variable y_i is the timber volume of the tree *i* (Vol_i); x_i is a vector of independent variables formed by the diameter to the height of the tree chest *i* (Dap_i) and the tree height (H_i). The model specified in loglinear form, known in the volumetric analysis literature as the Schumacher-Hall logarithmic model, is given by Equation 7.

$$lnVol_i = \beta_1 + \beta_2 lnDap_i + \beta_3 lnH_i + \varepsilon_i \quad (7)$$

The linear parameters found in the model can then be estimated by OLS.

On the other hand, when the derivative is a function of β, it is said that the model is nonlinear in the parameters, and may be specified in the form given by Equation 8.

$$Vol_i = \beta_1 Dap_i^{\beta_2} H_i^{\beta_3} + \varepsilon_i \quad (8)$$

It follows that, in this model, the derivatives depend on the β vector. Therefore, the OLS estimation method cannot be used to minimize the sum of squared residuals. In this case, the NLS method should be applied, which minimizes the sum of squared residuals with respect to the choice of β parameters. Thus, a better linear approximation of the population parameters can be obtained, with the estimation based on expanding the function $f(x, \beta)$ around the β_0 estimator, and applying the OLS to the final model. Throughout this iterative process of parameters estimation, in seeking the optimal solution, the first-order condition expressed in Equation 9 is satisfied.

$$[g(\beta)]'[y - f(x, \beta)] = 0 \quad (9)$$

Where $g(\beta)$ is the matrix of the first derivatives of $f(x, \beta)$ with respect to β. The expansion of the Taylor series for the optimal solution, in special cases, can be found in Pindyck and Rubenfield (2004).

Specification and adjustment of the volumetric model

One of the difficulties of multiple regression adjustment for estimation of the Amazon rainforest timber volume is the presence of atypical information. However, simply removing these observations can compromise the result. The characteristic of primary forests (native forests) to present commercial trees with higher-than-average value and with large volumetry.

To leave such observations in the sample dataset without adequately capturing their effects on the efficiency of the parameters, and with inadequate model specification, may incur violation of the homoscedasticity assumption (Santana, 2003); this tends to generate biased estimates of the average volume of wood.

To solve the problem, DVs are included in the equation to capture the effects of atypical values (values statistically positioned well above or well below the average of the data), due to the volume variation of the trees.

In this work, a high atypical observation (IA_i) is considered to be one that reaches a value greater than or equal to $IA_i = Q_3 + 3(Q_3 - Q_1)$, while a low observation is $(IB_i) = Q_1 - 3(Q_3 - Q_1)$. Here, Q_1 and Q_3 are, respectively, the first and third quartiles. Thus, the nonlinear model used in this study to estimate the volume of primary forest trees can be specified as in Equation 10:

$$Vol_i = \beta_1 Dap_i^{\beta_2} H_i^{\beta_3} e^{\beta_4 Vda_i + \beta_5 Vdb_i} + \varepsilon_i \quad (10)$$

Where Vol_i is the volume of the tree *i*, in m^3; Dap_i is the diameter of tree *i,* measured at breast height in m; H_i is the height of the tree bole *i* in m; Vda_i is the DV for an atypical high value of volume; Vdb_i is the DV for an atypical low value of volume well below the average; *e* is the base of the neperian logarithm; and ε_i is the error term normally distributed, independent and identically distributed $iid\ N(0, \sigma^2)$. The model specified in this way can only be estimated by NLS, generating nonbiased estimates and efficient parameters.

The nonlinear estimation process is determined by the iterative linearization method through the Taylor series expansion, which allows the function to be linearized around an initial set of values for the parameters, called β_0. In the sequence, the equation is estimated by OLS, generating new values for the parameters. In the second iteration, the nonlinear equation is linearized again around the new values estimated by OLS, and the process continues until it reaches convergence, or until the values of the parameters no longer change.

To illustrate this process, the Davidson and MacKinnon (2004) derivation is adopted. For this, the second-order expansion of the Taylor series is used, centred on the values of the vector β (Equation 11):

$$SQR(\beta) \approx SQR(\beta_0) + g'(\beta_0)(\beta - \beta_0) + \frac{1}{2}(\beta - \beta_0)H(\beta_0)(\beta - \beta_0) \quad (11)$$

Table 1. Volumetric models with and without DVs, estimated by NLS.

Dependent variable: *VOL= β_0*(DAP^β_1)*(HA^β_2)*exp(β_3*VDa+β_4*VDb)*				
Variable	**NLS with *dummy***		**NLS without *dummy***	
	Coefficient	**t-Statistic**	**Coefficient**	**t-Statistic**
Intercept - β_0	0.00012*	3.292598	6.09E-05*	7.18430
DAP - β_1	1.95425*	30.04372	2.09429*	78.8056
HA - β_2	0.79580*	23.19888	0.83039*	27.1666
VDa - β_3	0.08448*	2.837419		
VDb - β_4	-0.24604*	-4.929588		
R-squared	0.96387		0.96125	
Adjusted R-squared	0.96324		0.96091	
Sum squared resid	57.3815		61.5462	
Akaike info criterion	1.47501		1.52799	
Schwarz criterion	1.54884		1.57228	
Heteroskedasticity: Hw	0.952^{ns}		0.087^{ns}	
Multicollinearity: FVI	1.866^{ns}		1.015^{ns}	
Normality: Jb	4.812^{ns}		6.062**	
Durbin-Watson stat	2.132^{ns}		2.039^{ns}	
Chow Test F	8.3103 *	0.000327		

Source: Research data. *Significant at 1%; **Significant at 5%; ^{ns}not significant.

Where $g'(\beta_0)$ is the transpose matrix of the first differences of $f(x, \beta)$ with respect to β; and $H(\beta_0)$ represents the Hessian matrix (k x k) of $SQR(\beta)$, defined around β_0.
The first-order condition for a minimum is given by Equation 12.

$$g(\beta_0) + H(\beta_0)(\beta - \beta_0) = 0 \tag{12}$$

Which gives the origin to the iterative process, given by Equation 13.

$$\beta_{j+1} = \beta_j - \alpha H^{-1}(\beta_j)g(\beta_j) \tag{13}$$

Where α is a coefficient of adjustment, utilized in an iterative process, in order to improve the parameters' convergence to the true values. This is the nonlinear estimation process used by Stata12 software (2011) and Schwert (2009).

RESULTS

As shown in Table 1, the R² statistic indicates that there was a good fit of the model, in which 96.32% of the variation in the volume of trees was explained by the independent variables. Equally important is the observation that all independent variables had the expected signs and were statistically significant, based on the *t* statistic. The results agree with the theoretical assumptions for the commercial tree timber volume function of the Amazon forest, where the volume is a direct function of diameter and tree height.

The estimates of the parameters associated with these variables presented positive signs and significant values at 1%. The dependence relationship between these variables and the volume show that the diameter had a higher impact than did height in explaining changes in the volume of trees. Thus, for each 1% increment in the diameter value, the volume is expected to vary 1.95% in the same direction; therefore, variation in diameter produces a more than proportional influence on the volume of wood. Regarding height, each increase of 1% led to a 0.796% variation in the volume, that is, a less than proportional change.

The trees presented a volumetry that was well above or below the average of the sampled trees, which significantly influenced the estimated value of the volume of the trees' timber, as revealed by the significance of the *t*-statistics associated with the DVs. In this study, with the average volume of 4.43 m^3 as a reference point, the 13 atypical trees with a higher volume showed an average of 12.95 m^3, and an average of 1.65 m^3 for the six atypical trees with a smaller volume; that is, the values were 2.9 times higher and 2.7 times lower than the average, respectively.

The adapted Chow F test indicated the relevance of the two DVs included in the model; when absent, errors were found in the model specification, as was a bias in the volume estimates (Table 1). The Akaike and Schwarz criteria, by presenting lower statistical values for the model with DVs relative to the no-dummy model, also presented the best model specification. Finally, the Durbin–Watson statistic reflecting the result of the test for autocorrelation of residuals indicated the absence of this problem. This result was in line with expectations, since in cross-sectional samples for just one year this typical problem of time series data was not present.

The results of hypothesis tests for normality, hetero-

Table 2. Volumetric models with and without DVs estimated by OLS.

Dependent variable: *lnVOL*= $\beta_0 + \beta_1$ *lnDAP* + β_2 *lnHA* +β_3*VDa* + β_4*VDb*				
Variable	**OLS with *dummy***		**OLS without *dummy***	
	Coefficient	**t-Statistic**	**Coefficient**	**t-Statistic**
Intercept - β_0	-8.930925*	-40.74628	-9.508421*	-52.87361
lnDAP - β_1	1.905115*	40.46376	2.013940*	51.38402
lnHA - β_2	0.836789*	26.01590	0.878799*	26.72688
VDa - β_3	0.099373*	2.650997		
VDb - β_4	-0.225023*	-4.914108		
R-squared	0.948066		0.941702	
Adjusted R-squared	0.947159		0.941197	
Sum squared resid	2.502277		2.808908	
Akaike info criterion	-1.657508		-1.559007	
Schwarz criterion	-1.583676		-1.514708	
Heteroskedasticity: Hw	0.748^{ns}		0.952^{ns}	
Multicollinearity: FVI	1.629^{ns}		1.014^{ns}	
Normality: Jb	15.876*		18.639*	
Durbin-Watson stat	2.071^{ns}		1.933^{ns}	
Chow Test F	14.0311*	1.785E-06		

Source: Research data. *Significant to 1%; ^{ns}not significant.

scedasticity and multicollinearity confirmed that the model with the DV did not violate these assumptions, and that the restricted model without the incorporation of such variables displayed specification error, since the DVs were found relevant according to the significance of the test of normality and the *t*-statistics (Table 1).

From the results presented in Table 1, it can be observed that the atypical values (VDa and VDb) generated by trees influence significantly and in the opposite direction to the estimated volume of wood in logs. Additionally, the trees with lowest volume were found to cause a higher impact on estimates than the trees that produced larger atypical volumes, when compared to the average sample volume.

The results of the OLS for the logarithmic Hall–Schumacher model are shown in Table 2. It is noted that the DVs were relevant for specification of the logarithmic model, since the parameter estimates associated with such variables differed from zero at 1%. Likewise, the Chow F adapted test was also significant, ratifying the importance of including the variables in the equation. Moreover, the Akaike and Schwarz statistics showed lower values for the model with DVs, indicating a more adjusted model.

Regarding the assumptions of the regression model to be estimated by OLS, it was observed that the normality test was not met, demonstrating that R^2 statistics and regression standard deviation did not serve as parameters and analysis.

Finally, the estimated equations by NLS were compared with those estimated using OLS. The statistical adjusted R^2 of the equations estimated by NLS was higher by 2.3 and 2.01% for the equations with and without DVs, respectively. Given the improvement of accuracy through the model adjustment, and also due to the fact that the algorithm for estimating nonlinear equations is available in the primary software, it is worth making the effort to estimate the NLS equations.

DISCUSSION

To date, no study on volumetric equations of primary forests has referred to the treatment of atypical values of data samples through DVs. This gap also includes problems pertaining to violation of assumptions of the OLS heteroscedasticity model (or not) due to atypical values and multicollinearity were considered.

In this study, the Jarque and Bera (1987) test was applied to evaluate the normality of the residuals. The null hypothesis of normality of errors was accepted for the distribution of errors of the nonlinear model and rejected for the linear model estimated by OLS (Tables 1 and 2). This means that the Schumacher-Hall logarithmic model, which is considered as one having the best performance in volumetric studies (Souza and Jesus, 1991; Silva, 1984; Leite and Regazzi, 1992; Rolim et al., 2006; Thaines et al. 2010; Silva et al., 2011), should not be used without proper caution in the statistical evaluation, which directly influences the statistics and indicators used to select the “best fit” model. Indeed, the estimated NLS model met these assumptions, thereby validating the results.

Regarding the violation of the homoscedasticity hypothesis, Batista et al. (2004) and Silva et al. (2011) identify its infringement and correct the problems using weighted least squares and generalized least squares methods. In fact, some authors, such as Batista et al. (2004) and Silva et al. (2011), state that the variable volume tends to generate heteroscedasticity problems by causing different variance estimates in cross-sectional data.

Moreover, the disregard or unfamiliarity with the effects that violation of this hypothesis cause to the properties of estimators is found to be widespread. Indeed, many studies have included graphics between the error terms and the independent variables that show a distribution in the form of a funnel, as in Thaines et al. (2010), thereby indicating that the variance of the errors is not constant.

Furthermore, Rolim et al. (2006) identify a growing trend among the regression errors and independent variables that reveals the presence of heteroscedasticity. However, it is surprising that nothing is mentioned about correcting the problem, with results published and faults going unnoticed in published journals. This absence of care in evaluating the regression analysis assumptions is a serious issue related to abandon the rigor that should be given to the use of statistical tools and the disservice this causes to those who use these results to make decisions.

The multicollinearity hypothesis is tested by Silva et al. (2011). In other studies (Batista et al., 2004; Rolim et al., 2006), this hypothesis is only mentioned in passing, at most, which also reflects the lack of concern regarding the high correlation between the independent variables. This ratifies the importance of understanding the many problems that multicollinearity causes, including the instability of the estimated parameters, which directly affects projections of estimated values of the dependent variable. In fact, polynomial models that use the squared variables and their combinations make the partial and/or multiple correlations between them very high and generally transformed, presenting a high level of multicollinearity. As a result, though the statistical R^2 presents high value, significant parameters are not found and signals are exchanged (Santana, 2003; Pindyck and Rubinfeld, 2004; Gujarati, 2003). Due to this problem, and especially the nonhomogeneity of the residuals, these models are rejected in favour of the of Schumacher-Hall logarithmic model, which does not violate these assumptions.

With respect to the use of DVs, some studies apply this technique to explain the influence of different regions on estimating the average volume of wood (Mc Tague et al., 1989; Batista et al., 2004). Nevertheless, no record has been found that applies DVs to capture the effects of atypical information on estimation of the average volume of wood.

However, in the case of samples of primary forests, especially the Amazon forest, atypical observations provide desirable information because they represent species of high commercial value and, for the most part, are regarded as the most valuable data observations. As these observations exert desirable influence on the parameters, they must be directly incorporated into the model specification using dummies in order to test their relevance, regardless of the form of the volumetric equation. This study applied this technique in a pioneering way and the results were significant for the adjusted NLS and OLS models (Tables 1 and 2). The adapted Chow F test, together with the Akaike and Schwarz criteria, confirmed that the atypical observations were well specified in both models via the use of DVs. Thus, in addition to eliminating the specification error of the models, we were able to incorporate a significant influence of these observations. In the case of primary forests, it is important for future studies to observe the analysis of atypical observations and their specification for the evaluation of normality assumptions, the estimation of NLS models, and heteroscedasticity and multicollinearity, in the case of cross-sectional data.

With respect to the estimation method, some studies have applied the Schumacher-Hall general model or nonlinear model (Mc Tague et al., 1989; Batista et al., 2004) via NLS and concluded that some models show better adjustment. In addition, it is not necessary to correct the variables used to estimate volume. The difference between this work in relation to others is the use of the NLS generalized method to estimate the correction of heteroscedasticity using White's method, as it generates consistent results for the variance and covariance matrix, and provides asymptotically valid statistical tests, while the weighted methods used in other studies do not guarantee a good adjustment for the model (Pindyck and Rubinfeld, 2004).

Finally, it is emphasized that, following the adjustment made via application of the Furnival Index, the comparison between the standard errors of the NLS (0.50057) and OLS (0.73787) estimated models showed the former to have a standard error that was 32.16% lower than that of the OLS. The comparison via adjusted R^2 between the NLS (adjusted R^2 = 0.9632) and OLS (adjusted R^2 = 0.9287) models, after correction for comparison purposes (Gujarati, 2003), allowed us to demonstrate that the nonlinear model showed a 3.71% superiority in the adjustment. These results confirm that the Schumacher-Hall nonlinear model allows for better adjustment compared to the logarithmic model.

Conclusion

The inclusion of dummy variables to capture the effects of relevant atypical observations made the specification of models more appropriate and generated better estimation and adjustment of parameters using OLS and NLS. Failure to incorporate the effects of atypical observations in specification of the models necessarily implies a specification error and over- or underestimation of the timber volume for the company. In addition, it distorts economic results, such as profit, and the productivity factors of labour and capital. Due to the good performance of the wood volumetric model for trees grown in primary forests, in which atypical values were represented by DVs and estimation was performed using NLS, it is suggested that the model be tested in cases of Amazon's public areas, which are intended as forest concessions by the Brazilian Forest Service, and also in Federal and State public areas, particularly the protected areas selected for sustainable use, such as the National Forests (Flonas) and State Forests (Flotas).

Conflicts of Interests

The authors have not declared any conflict of interests.

REFERENCES

Batista JLF, Marquesini M, Viana VM (2004). Equações de volume para árvores de Caxeta (Tabebuia cassinoides) no Estado de São Paulo e sul do Estado do Rio de Janeiro. Scientia Forestalis Piracicaba 65:162-175.

Chow GC (1960). Tests of equality between sets of coefficients in two linear regressions. Econometrica: J. Econ. Society pp. 591-605.

Davidson R, MacKinnon JG (2004). Econometric theory and methods (Vol. 5). New York: Oxford University Press.

Schwert G (2009). Eviews 7 user's guide II. Quantitative Micro Software, LLC, Irvine, California.

Furnival GM (1961). An index for comparing equations used in constructing volume tables. Forest Sci. *7*(4):337-341.

Gujarati DN (2003). Basic Econometrics. 4th. New York: McGraw-Hill. 1032p.

Jarque CM, Bera AK (1987). A test for normality of observations and regression residuals. Int. Stat. Review/Revue Int. de Statistique. pp. 163-172.

Leite HG, Regazzi AJ (1992). Métodos estatísticos para avaliar a igualdade de equações volumétricas. Rev. Árvore 16(1):59-71.

Mc Tague JP, Batista JLF, Steiner L (1989). Equações de volume total, volume comercial e forma do tronco para plantações de Eucalyptus nos estados de São Paulo e Rio de Janeiro. IPEF Piracicaba (41-42):56-63.

Pindyck RS, Rubinfeld DL (2004). Econometria: modelos & previsões. Elsevier. 726p.

Rolim SG, Couto HD, Jesus RD, França JT (2006). Modelos volumétricos para a Floresta Nacional do Tapirapé-Aquirí, Serra dos Carajás (PA). Acta Amazonica, Manaus 36(1):107-114.

Santana AC (2003). Métodos quantitativos em economia: elementos e aplicações. Belém: Ufra P 485.

Santana AC, Santos MAS, Santana ÁL, Yared JAG (2012). O valor econômico da extração manejada de madeira no Baixo Amazonas, Estado do Pará. Rev. Árvore 36(3):527-536.

Santana AC, Amin MM, Costa NL, Santana ÁL, Santos MAS (2016). An economic valuation of managed roundwood in the region of marajó, in the state of Pará. Int. J. Dev. Res. 6(3):7222-7227.

Santos PC, Santana AC, Barros PLC, Queiroz JCB, Oliveira T (2011). O emprego da geoestatística na determinação do tamanho "ótimo" de amostras aleatórias com vistas à obtenção de estimativas dos volumes dos fustes de espécies florestais em Paragominas, estado do Pará.

Schumacher F, Hall F (1933). Logarithmic expression of timber-tree volume. J. Agric. Res. 47(9):719-734.

Silva JNM, Carvalho JOP, Lopes JCA, Carvalho MSP (1984). Equações de volume para a floresta nacional do tapajós. In *Boletim de* Pesquisa Florestal Numero 8/9. Empresa Brasiliera de Pesquisa Agropecúaria, Centro de Pesquisa Agropecúaria do Tropicos Umidos Belém, Pará.

Silva EN, Santana AC, Queiroz VT, Souza RJ (2011). Estimação de equações volumétricas para árvores de valor comercial em Paragominas, estado do Pará. Amazônia: Ciênc. Desenvolvimento 7(2):7-18.

Souza AL, Jesus RM (1991). Equações de volume comercial e fator de forma para espécies da Mata Atlântica ocorrentes na reserva florestal da Companhia Vale do Rio Doce, Linhares, ES. Rev. Árvore 15(3):257-273.

Stata12 (2011). Stata base referencial manual volume 3*.* Texas: StataCorp LP pp. 1253-1914.

Thaines F, Braz EM, Mattos PP, Thaines AAR (2010). Equações para estimativa de volume de madeira para a região da bacia do rio Ituxi, Lábrea, AM. Pesquisa Florestal Bras. Colombo 30(64):283-289.

White H (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica: J. Econ. Society pp. 817-838.