Research Journal of

Forestry

ISSN 1819-3439

Academic Journals Inc.

www.academicjournals.com

Research Journal of Forestry 5 (2): 50-65, 2011
ISSN 1819-3439 / DOI: 10.3923/rjf.2011.50.65

Improved Stem Volume Estimation using P-Value Approach in Polynomial Regression Models

Noraini Abdullah, Zainodin Hj. Jubok and Amran Ahmed
School of Science and Technology, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia

Corresponding Author: Noraini Abdullah, School of Science and Technology, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia

ABSTRACT

Sustainable forestation, its management and practices seek the importance for an estimation tool. Hence, the aim of this study was to estimate the parameters of the dependent variable (stem volume) having a polynomial relationship with the independent variables using the Polynomial Regression (PR) technique. Field data collection involved measurements on 130 trees of an indigenous timber species, *Cinnamomum iners*. The stem height, diameter at base, middle and the top of the stem before the crown were tree measurements taken as variables. The stem volume was based on the Newton's formula. Variables normality and linearity exhibited polynomial characterization of power terms greater than 2. Thirty-two Polynomial Regression Models (PRM) with the auxiliary variables were considered up to their third order interactions. Preliminary, multicollinearity between the independent variables was minimized and statistical tests involving the Global, Coefficient and Wald tests were carried out to select significant variables with their possible interactions. Comparisons between the Polynomial Regression Models (PRM) were made using the eight selection criteria (8SC). The best regression model (P26.5.3) with five multicollinearity and three insignificant variables removed, was identified based on the minimum value on majority of the 8SC. The Goodness-of-fit tests were done to validate the chosen best model. The use of an appropriate transformation was found to have increase in the degree of a statistically valid polynomial, hence, providing an improved estimation for tree stem volume.

Key words: Polynomial regression models (PRM), normality, linearity, multicollinearity, eight selection criteria (8SC)

INTRODUCTION

Research information on tree stem volume estimation with nondestructive felling methods was few. Hence, the vitality of an integrated approach in identifying factors affecting tree growth. Revenues from agro-forestry and sustainable forest products were some of the economic gains benefited from proper modeling procedures adopted and appropriate models chosen for estimation and prediction. These parameterization methods can also be utilized in varying areas, such as in forest ecology where carbon assessment and sequestration are optimized (Cao *et al.*, 2010) and in ecological modeling of stem biomass accumulation (Kubo and Kobayashi, 2007). As an example, John (2003) wrote on models for estimation and simulation of crown and canopy cover, covering eight states in the United States. He used measurements of the diameter at breast height, crown widths, crown ratio, number of stems, density measurements, elevation and various geographical or political divisions as variables in his models. Shawn *et al.* (2007) had stated that tree crowns are the sources of carbohydrates from photosynthesis, hence, serving as a vigor indicator of the tree. The stem (or bole) diameter was also found to be statistically significant and the strongest predictor in all models of tree estimations and for some species a quadratic term was needed for model enhancement (Bechtold, 2004). Volume estimation using non-linear regression models by Adekunle (2007) in the natural forest ecosystem in Southwest Nigeria had replaced tree age with diameter at breast height or basal area in their original model and used the standard error of estimate for model fitness. Alderete-Chavez *et al.* (2010) had evaluated tree growth for reforestation by measuring tree height and diameter while Peter and Oluwafeni (2008) had also used stem form and tree size as their independent variables in their works of testing four non-linear individual tree functions for interim crown ratio predictions. The importance of stem-form in describing the tree crown was reflected in the Logistic and Richards functions. Hasenauer (2006) had used modelling to predict future forest stand development, Noraini *et al.* (2008) had presented tree stem biomass prediction and estimation with Newton's and Huber's functions using multiple regression models. They had employed the backward elimination method to remove insignificant variables from their models. The resultant best equation presented a simplified model for log volume estimation.

By modelling tree stem volumes, other benefits can be gained attributed from its models and estimation. As an example, the green canopy cover of the urban forest could be estimated and indirectly, the photosynthetic functions of tree crowns and gas cycles in urban areas could be monitored and ascertained its environmental qualities. Hence, in this paper, a modeling approach was incorporated with the employment of the Polynomial Regression (PR) technique to estimate the parameters of the volumetric models whereby factors affecting tree growth were considered.

MATERIALS AND METHODS

Study area: The study was done on an indigenous tree species, *Cinnamomum iners*, planted as shade trees alongside roads and in parks. However, the field data measurements were collected in Universiti Malaysia Sabah and around the city of Kota Kinabalu. The university is located on a 999 acres piece of land along Sepanggar Bay, about 5 km north-east of the city of Kota Kinabalu. It is situated at latitude 6°00'N and longitude 116°04'E (http://www.mapsofworld.com/) in a tropical rainforest zone with a mean annual rainfall of around 2000-2499 mm and a relative humidity of 81.2±0.3°C. The annual temperature is 27±0.1°C and its rainy season extends from October to April (http://www.met.gov.my).

Stem biomass volumetric equation: During field data collection in 2009, 130 trees were measured non-destructively and randomly selected from five sample plots in the vicinity of Kota Kinabalu. The materials used in field data measurements would include a clinometer and a girth tape. The clinometer was used to measure the height of a tree while the fiberglass girth tape was used to measure the diameter indirectly. This was done by firstly wrapping it round the tree to measure the circumference in a perpendicular plane to the stem axis and then its value was divided by PI (π) to estimate the diameter. The tree variables considered were diameter at the base (D_b), diameter at the middle (D_m), diameter at the top (D_t) and the stem height (T) (Husch *et al.*, 2003; Avery and Burkhart, 1994; Philip, 1994). The cylindrical area of log (A) was given by formula πR^2, where R is the radius of the tree stem or bole. Hence, at each respective sections of the log, the area will be known as A_b (at the base), A_m (at the middle) and A_t (at the top). Since diameter is twice the radius, the corresponding cross-sectional area will then be calculated as: $A_b = \pi D_b{}^2/4$, $A_m = \pi D_m{}^2/4$ and $A_t = \pi D_t{}^2/4$, respectively.

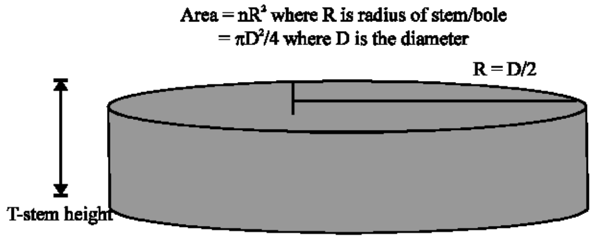

Fig. 1: Merchantable tree log

The volume of the merchantable log would be the stem height multiplied by the area as shown in Fig. 1. In this study, based on these mensuration variables, the volume of the stem biomass was calculated using the Newton's formula as in study of Fuwape *et al.* (2001). The objective of this study was to compare models based on the Newton's Formula using the Polynomial Regressions (PR) technique with significant attribution to the power terms.

Model-building description

Data preparation and the p-value approach for normality: In regression analysis, normality and linear relationship of data is of prime importance and this is unachievable with most environmental and forestry data. Hence, normality test was initially carried out numerically (Coakes and Steed, 2007), with the graphical method as supporting evidence (Ashish and Muni, 1990). Normality tests were carried out in SPSS and the test statistic of the variable was given by the statistic value of Kolmogorov-Smirnov (for $n>50$) and Shapiro Wilk (for $n<50$). The confidence level was set at 95% which is the standard percentage of the normality test with $\alpha>5\%$ as significant.

The type of transformation would be identified by first plotting a scatter plot of the dependent variable over the independent variable. For normality and linearity, appropriate transformation values were taken to form the characterization of the polynomial terms. Driving up or driving down the ladder would depend on the concavity or convexity of the scatter plot. This study would focus on the searching efforts for the best transformations applicable to the data sets and optimizes the range of transformation needed. These would become a part of models' simulations and optimization approach in getting the best model.

Transformation would involve: (1) identifying the types of curvilinear data, (2) determining the types of transformation needed and (3) exercise the procedures for Ladder, Power or Box-Cox transformations. The power transformation procedure used the data sets with the power of the origin being employed which was given by Devore and Peck (1993).

Using the p-value from the F-statistics, data with p-values >0.05 were considered as normal. Several iterations were executed so to determine the best transformation required for normality. Figure 2 depicted the flowchart on the data transformation procedures using Ladder and Box-Cox transformations executed on non-normal or nonlinear data before any model building could be developed.

Modelling and model-building approach: Figure 3 depicted the modelling flowchart. Preliminary with the conceptual development of the importance of modelling, its estimations and contributions to the real world problems, mathematical theories were applied for model building.

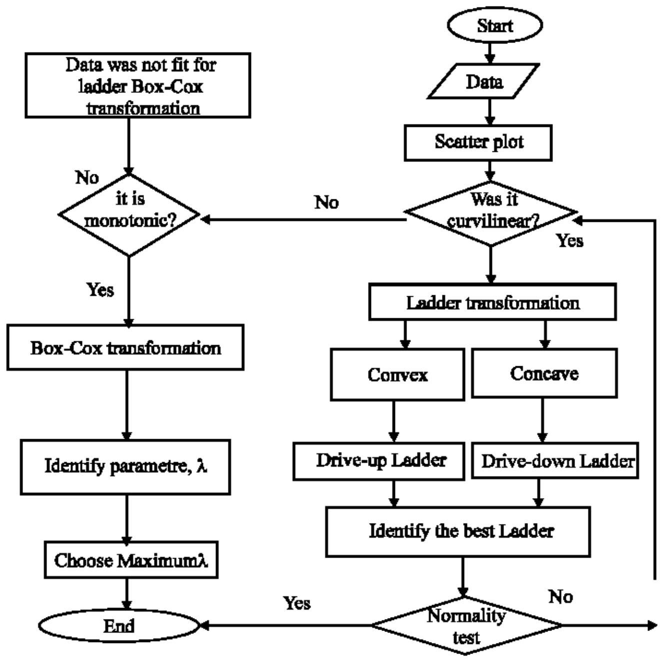

Fig. 2: Flow chart on the procedures of data transformations

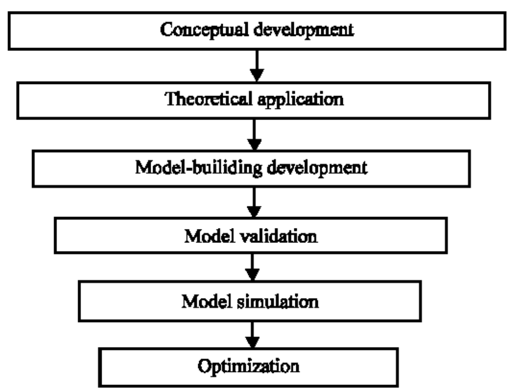

Fig. 3: Modelling flowchart

Figure 4 showed the four phases of the Model-building development. Model-building techniques were exemplified and validated through tests and hypotheses. Model's validation was enhanced by simulation and optimization of values, expected to be characterized as optimal values.

Polynomial Regression Models (PRM): Phase 1 of Model-building in Fig. 4, consisted of the all possible models which were made up of variables that had been prepared after undergoing the data preparation procedures of Fig. 2. For simplicity, these variables were then known as the defined transformed variables.

The PR models were made of a dependent variable, V, the stem volume and single independent variables, taken from field data mensuration. The model-building was developed based on the method of multiple regressions, a statistical method of more than two independent variables as in Eq. 1:

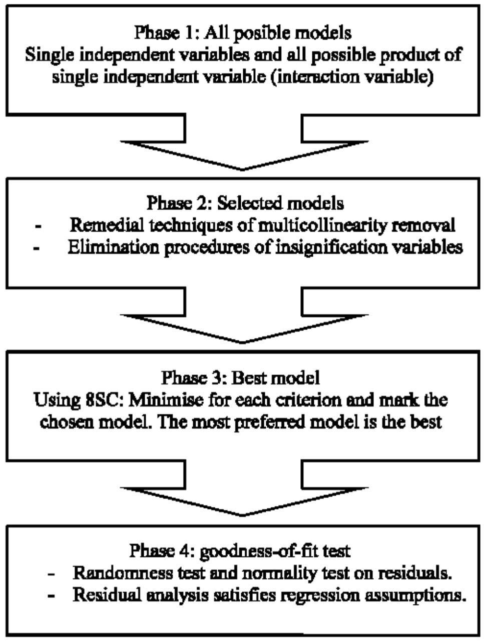

Fig. 4: The four phases in model-building development

$$Y_i = \beta_o+\beta_1X_{1i}+\beta_2X_{2i}+\cdots+\beta_kX_{ki}+a_i \quad (1)$$

where, i=1, 2, ... , n; Y_i was the dependent variable; X_{1i}, X_{2i}, ... , X_{ki} were the independent variables; β_i's were the regression coefficients with k+1 parameters and a_i's were the error terms. As with polynomials of the order 2 (parabolic curve with quadratic terms), the model equation could be written as:

$$Y_i = \beta_o+\beta_iX_{1i}+\beta_{11}X_{ii}^2+\cdots+\beta_kX_{ki}+\beta_{kk}X_{ki}^2+a_i \quad (2)$$

The number of models as given by the formula: $N=\sum_{j=1}^{q} j({}^qC_j)$, where 'N' represented the number of possible models, 'q' represented the total number of single independent variables for j=1, 2, 3,... , q. Based on say four single independent variables, the number of models would be the sum of the combinations of all the four variables which would be, $(1x^4C_1)+(2x^4C_2)+(3X^4C_3)+(4x^4C_4)$, thus, 32 models, as shown in Table 1.

Examples of possible PRM's were shown in Table 2 whereby models from P1-P15 were without interactions, P16-P26 (1st order interactions), P27-P31 (2nd order interactions) and P32 (3rd order interactions). The all possible PR models were listed as in the Appendix.

One of the possible models with different variables' attributes was given by model P27:

$$V_{27} = \beta_o+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_{11}X^2_1+\beta_{22}X^2_2+\beta_{33}X^2_3+\beta_{12}X_{12}+\beta_{13}X_{13}+\beta_{23}X_{23}+\beta_{123}X_{123}+a_{27} \quad (3)$$

with X_1, X_2 and X_3 as the single independent variables, X_{12}, X_{13} and X_{23} as the 1st order interactions, X_{123} as the 2nd order interaction and X_1^2, X_2^2 and X_3^2 as the polynomial terms of power 2 (or also known as the quadratic terms) and a_{27} was the residual term for the model P27.

Table 1: Total number of possible models

No. of variables	Single independent variables	Order of interactions 1st	2nd	3rd	Total No.of models
1	4	-	-	-	4
2	6	6	-	-	12
3	4	4	4	-	12
4	1	1	1	1	4
Total	15	11	5	1	32

Table 2: All possible models of four single independent variables

P1	$V_1 = \beta_0+\beta_1X_1+\beta_{11}X_1^2+a_1$
P2	$V_2 = \beta_0+\beta_2X_2+\beta_{22}X_2^2+a_2$
P3	$V_3 = \beta_0+\beta_3X_3+\beta_{33}X_3^2+a_3$
P4	$V_4 = \beta_0+\beta_4X_4+\beta_{44}X_4^2+a_1$
:	: : :
P15	$V_{15} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+a_{15}$
P16	$V_{16} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{12}X_{12}+a_{16}$
:	: : : : : :
P27	$V_{27} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{12}X_{12}+\beta_{13}X_{13}+\beta_{23}X_{23}+\beta_{123}X_{123}+a_{27}$
:	: : : : : :
P32	$V_{37} = \beta_0+\beta_1X_1+\cdots+\beta_{11}X_1^2+\cdots+\beta_{12}X_{12}+\cdots+\beta_{34}X_{34}+\beta_{123}X_{123}+\cdots+\beta_{1234}X_{1234}+a_{32}$

The models could then be written in a general form as:

$$V_{PR} = \Omega_0+\Omega_1W_1+\Omega_2W_2+\cdots+\Omega_kW_k+u \quad (4)$$

where, V_{PR} was the volume, '*W*' was an independent variable which could represent one of these types of variables, namely, single independent, interactive, generated, transformed, quadratic terms or even dummy variables, Ω's were the newly defined regression coefficients and 'u' as the error terms for each respective transformed model.

Multicollinearity removal and insignificant variable elimination: Multicollinearity was a statistical phenomenon where very strong linear or perfect relationships would exist between the independent variables, thus affecting the Sum of Square Error (SSE) of the respective models (Gujarati, 2006) and colhnearity between the variables could be identified by examining the values of the correlation matrix of the independent variables. High correlation coefficients of absolute values in the range of $0.75\geq|r|\geq0.95$ were considered to exhibit multicollinearity effects. These effects of multicollinearity could be remedied by doing model transformations or by removing the highly correlated source variables. These multicollinearity source variables had to be dealt with first before modelling could be done, as indicated in Phase 2 of model development in Fig. 4. Absolute coefficient values that, were greater than 0.75, were normally considered to cause multicollinearity effects. However, in this paper, high correlation coefficient of absolute values greater than 0.95 ($|r|\geq0.95$) were chosen to be removed.

A case I type of multicollinearity problem occurred when there was a presence of a high frequency with no tie in the correlation coefficient matrix. This independent variable having the highest absolute $|r|\geq0.95$ would be removed and the model was thus rerun to check any more occurrences of multicollinearity.

Types
Description
Procedures

Case I

Highest Frequency with no tie
-An independent variable has high |r|=0.95 among variables. Remove corresponding variable and model is rerun.

Case II

Highest Frequency and tie found
-Some independent variables have same frequency with high |r|=0.95. Remove the weakest/least value with the dependent variable and rerun model.

Cases III

Highest Frequency with only one tie
-Two pairs of mnlticollinearity variable have the same frequency of 1. Remove the weakest/least with the dependent variable and rerun model.

Fig. 5: Multicollinearity Removal Procedures

Contrarily, a case II type problem occurred when there was a presence of some independent variables of the highest frequencies with a tie found in the correlation matrix. The variable with the highest frequency and the least value or the weakest contribution to the dependent variable would be removed. The model was then rerun to check any more occurrences of multicollinearity.

A case III type would involve two pairs of multicollinearity source variables of the highest frequency with only a single tie. By comparing these two independent variables, the variable with least value or the weakest contribution to the dependent variable was removed. The model was then again rerun to check the presence of multicollinearity source variables.

All the models might exhibit any one of these cases, or a combination of these case-types in their correlation coefficient matrices. The remedial techniques of multicollinearity removal were executed in Phase 2 of the model-building development. Figure 5 showed the Case Types for Multicollinearity removal procedures that would be carried out on the all possible models.

The elimination of insignificant variables from the models (or Coefficient Test) was carried out using the backward elimination method. Illustrations of the backward elimination method had been shown by Noraini *et al*. (2008).

Best model selection and goodness-of-fit tests: Table 3 depicted the selection of the eight criteria (8SC) of Phase 3, used in identifying the best regression model (Ramanathan, 2002). The criteria were based on the value of Sum of Square Error (SSE) where n would be the number of samples or observations and (k+1) was the number of parameters in each respective model. The model having the least value in majority of the criteria would be chosen as the best model.

The best model would undergo the Goodness-of-Fit tests of Phase 4 in Fig. 4 which comprised of the normality and randomness tests on the models' residuals. Without violating the assumptions in regression analysis, further simulations of the best model would provide a better prediction for stem volume estimation in future forest planning strategy and management.

Table 3: Eight Selection Criteria (8SC) for best model identification

Criteria	Model	Criteria	Model
AIC (Akaike, 1974)	$\left(\frac{SSE}{n}\right)e^{(2(k+1)/n)}$	RICE (Rice, 1984)	$\left(\frac{SSE}{n}\right)\left[1-\left(\frac{2(k+1)}{n}\right)\right]^{-1}$
FPE (Akaike,1970)	$\left(\frac{SSE}{n}\right)\frac{n+(k+1)}{n-(k+1)}$	SCHWARZ (Schwarz, 1978)	$\left(\frac{SSE}{n}\right)n^{(k+1)/n}$
GCV (Golub *et al*., 1979)	$\left(\frac{SSE}{n}\right)\left[1-\left(\frac{k+1}{n}\right)\right]^{-2}$	SGMASQ (Ramanathan, 2002)	$\left(\frac{SSE}{n}\right)\left[1-\frac{(k+1)}{n}\right]^{-1}$
HQ (Hannan and Quinn, 1979)	$\left(\frac{SSE}{n}\right)\ln n^{2(k+1)/n}$	SHIBATA (Shibata, 1981)	$\left(\frac{SSE}{n}\right)\frac{n+2(k+1)}{n}$

Where, n would be the number of observations, (k+1): the number of model's parameters and SSE the sum of square of error. AIC : Akaike Information Criterion, RICE: Rice, FPE: Finite Prediction Error, SCHWARZ: Schwarz Criterion, GCV: Generalised Cross Validation, SGMASQ: Ramanathan, HQ: Hannan and Quinn Criterion, SHIBATA: Shibata

RESULTS AND MODEL ANALYSIS

Transformed variables for normality and linearity: Normality tests and Power transformations were carried out on the non-normal data. Table 4 depicted the defined variables, before and after transformations.

From Table 5, the p-values of variable D_t increased in the variable power range of 1.5-3.5, before decreased to the value of 4.5. The optimal (highest) p-value was 0.034 and the variable power was thus focused at 3.5 where 3 was the first digit of transformation. The second transformation digit would thus between the values of 3.5-4.5.

Referring to Table 6, variable D_t had reached the optimal normality p-value of 0.061(highest) at the transformation value of 3.7. The third transformation digit which was the second decimal digit would lie between 3.65-3.75. Similar procedures were executed on the other variable, D_m. A generated variable, D_b/T, had been created for normality. Table 7 depicted the descriptive statistics of the models' transformed variables. All the transformed variables had turned to normal since the significant p-value were more than 0.05. The data sets could then be used for further regression analysis.

Multicollinearity removal and backward elimination method: After transformation, then followed by Phase 2 of the selected model whereby two procedures were involved, viz., the remedial techniques of multicollinearity removal and the elimination procedures of the insignificant variables. Parameter tests were also carried out during the model selection. This would include the Global (or F-) Test where the overall significance of the independent variables on the dependent variable was tested, the Coefficient Test where the backward elimination procedures were implemented and the Wald test which was to test the overall goodness-of-fit of the stem volume regression model, by checking the effect of the omitted independent variables on the volumetric stem biomass, V.

The multicollinearity removal procedures employed in Phase 2 would not be dealt with in detail but then sufficed to include the coefficient correlation matrix of the best model before and after multicollinearity removal and elimination of insignificant variables being carried out (Table 8-10), respectively. The highlighted values (colored grey) in Table 8 indicated examples of high correlation values exhibiting multicollinearity effects of Case II type of the independent variables (X_1, X_2, X_{12}).

Table 4: Definition of variables before and after transformation

Variable	Definition	Transformation	Transformed variables
V_{nw}	Stem Volume.(m^3): Nw-Newton	V_{nw}	V_1
D_t	Diameter at top of trunk	$D_t^{3.7}$	X_1
D_m	Diameter at middle of trunk	$D_m^{4.5}$	X_2
D_b	Diameter at the base of trunk	D_b/T	X_3
T	Tree height (m)	T	X_4

Table 5: Normality test using Kolmogorov Smirnov on variable D_t

	Kolmogorov-Smirnov		
Transformed variable	Statistics	df	p-value
$D_t^{1.5}$	0.148	130	0.000
$D_t^{2.5}$	0.115	130	0.000
$D_t^{3.5}$	0.082	130	0.034
$D_t^{4.5}$	0.090	130	0.011

Table 6: Normality Test on Focus Optimal value of Variable D_t

	Kolmogorov-Smirnov		
Transformed variable	Statistics	df	p-value
$D_t^{3.6}$	0.078	130	0.049
$D_t^{3.7}$	0.076	130	0.061
$D_t^{3.8}$	0.078	130	0.051
$D_t^{3.9}$	0.080	130	0.043

Table 7: Descriptive statistics of transformed variables

	Transformed variables				
Defined variables	V	X_1	X_2	X_3	X_4
Mean	0.9215	0.1360	0.1081	0.1070	6.1303
Variance	0.133	0.005	0.004	0.000	0.896
Std. Deviation	0.3643	0.0713	0.0628	0.0144	0.9466
Minimum	0.180	0.010	0.00	0.070	3.780
Maximum	1.960	0.400	0.330	0.150	8.230
Skewness	-0.020	0.331	0.332	0.624	-0.257
Kurtosis	-0.147	0.602	0.158	0.905	-0.378
Kolmogorov-Smirnov	0.068	0.076	0.060	0.065	0.043
Kolmogorov-Smirnov (sig. p-value)	0.200	0.061	0.200	0.200	0.200

Standard error (s.e.) of Skewness is 0.212. Standard error (s.e.) of kurtosis is 0.422

They exhibited a tie of frequency 2. However, variable X_{12} had the weakest value (0.807) on the dependent variable which then resulted in the first multicollinearity removal of variable X_{12}. The model was then rerun to become model P26.1 with digit of value 1 as the first multicollinearity removal.

Table 9 showed the removal of the second multicollinearity source variables of Type Case II, indicated by variables (X_1, X_2). X_2 was removed since it had the least value (0.884) on the dependent variable, followed by the removal of variable X_1. The resultant model then became model P26.3 where digit of value 3 indicated three multicollinearity removals.

Table 8: Correlation coefficient matrix of model p26.0

	V	X_1	X_2	X_3	X_4	X_{12}	X_{13}	X_{14}	X_{23}	X_{24}	X_{34}	X_{11}	X_{22}	X_{33}	X_{44}
Y	1														
X1	0.897	1													
X2	0.884	0.917	1												
X3	0.116	0.219	0.225	1											
X4	0.859	0.641	0.619	-0.332	1										
X12	0.807	0.904	0.922	0.224	0.522	1									
X13	0.841	0.969	0.896	0.413	0.503	0.908	1								
X14	0.940	0.979	0.904	0.101	0.747	0.904	0.919	1							
X23	0.834	0.901	0.975	0.393	0.497	0.927	0.934	0.861	1						
X24	0.924	0.908	0.982	0.117	0.714	0.924	0.859	0.931	0.932	1					
X34	0.905	0.786	0.770	0.459	0.677	0.689	0.814	0.791	0.794	0.773	1				
X11	0.789	0.936	0.838	0.211	0.516	0.954	0.931	0.929	0.848	0.844	0.673	1			
X22	0.783	0.849	0.937	0.210	0.505	0.981	0.853	0.853	0.934	0.937	0.662	0.888	1		
X33	0.095	0.200	0.202	0.995	-0.342	0.209	0.399	0.082	0.374	0.094	0.441	0.199	0.194	1	
X44	0.855	0.626	0.604	-0.334	0.995	0.518	0.485	0.742	0.479	0.709	0.666	0.512	0.501	-0.341	1

Table 9: Correlation coefficient matrix of model P26.1

	V	X_1	X_2	X_3	X_4	X_{13}	X_{14}	X_{23}	X_{24}	X_{34}	X_{11}	X_{22}	X_{33}	X_{44}
Y	1													
X1	0.897	1												
X2	0.884	0.917	1											
X3	0.116	0.219	0.225	1										
X4	0.859	0.641	0.619	-0.332	1									
X13	0.841	0.969	0.896	0.413	0.503	1								
X14	0.940	0.979	0.904	0.101	0.747	0.919	1							
X23	0.834	0.901	0.975	0.393	0.497	0.934	0.861	1						
X24	0.924	0.908	0.982	0.117	0.714	0.859	0.931	0.932	1					
X34	0.905	0.786	0.770	0.459	0.677	0.814	0.791	0.794	0.773	1				
X11	0.789	0.936	0.838	0.211	0.516	0.931	0.929	0.848	0.844	0.673	1			
X22	0.783	0.849	0.937	0.210	0.505	0.853	0.853	0.934	0.937	0.662	0.888	1		
X33	0.095	0.200	0.202	0.995	-0.342	0.399	0.082	0.374	0.094	0.441	0.199	0.194	1	
X44	0.855	0.626	0.604	-0.334	0.995	0.485	0.742	0.479	0.709	0.666	0.512	0.501	-0.341	1

Table 10: Correlation coefficient matrix of model P26.3

	V	X_3	X_4	X_{13}	X_{14}	X_{23}	X_{24}	X_{34}	X_{11}	X_{22}	X_{33}	X_{44}
Y	1											
X3	0.116	1										
X4	0.859	-0.332	1									
X13	0.841	0.413	0.503	1								
X14	0.940	0.101	0.747	0.919	1							
X23	0.834	0.393	0.497	0.934	0.861	1						
X24	0.924	0.117	0.714	0.859	0.931	0.932	1					
X34	0.905	0.459	0.677	0.814	0.791	0.794	0.773	1				
X11	0.789	0.211	0.516	0.931	0.929	0.848	0.844	0.673	1			
X22	0.783	0.210	0.505	0.853	0.853	0.934	0.937	0.662	0.888	1		
X33	0.095	0.995	-0.342	0.399	0.082	0.374	0.094	0.441	0.199	0.194	1	
X44	0.855	-0.334	0.995	0.485	0.742	0.479	0.709	0.666	0.512	0.501	-0.341	1

Table 10 showed the multicollinearity variables of Case III and Case I types on variables X_{33} and X_{44}, respectively. X_{33} was initially removed since it had the least value (0.095) on the dependent variable which was then followed by the removal of variable X_{44}.

Subsequent five multicollinearity source variables had been removed thus resulted in the correlation coefficient matrix of model P26.5.0 as shown in Table 11. Table 11 showed that inexistence of high multicollinearity variables in the model. The procedures of eliminating insignificant variables could therefore be carried out using the backward elimination. It could be seen that variables (X_4, X_{13} and X_{11}) were removed since they were not significant. Table 12 depicted the final matrix whereby all the remaining variables in the model were significant with their p-values of more than 0.05 ($\alpha \geq 5\%$).

Comparisons of best model regression: The best model from the 8SC was based on the (k+1) parameters and fulfilled the least value of most of the criteria (Ramanathan, 2002). Table 13 signified the comparisons of the PR models based on the eight selection criteria. It could be seen that the best PR model was represented by the model P26.5.3 with five multicollinearity removals and three insignificant variables eliminated.

Table 11: Correlation coefficient matrix of model P26.5.0 without multicollinearity

	V	X_3	X_4	X_{13}	X_{14}	X_{23}	X_{24}	X_{34}	X_{11}	X_{22}
V	1									
X_3	0.115	1								
X_4	0.858	-0.331	1							
X_{13}	0.840	0.412	0.502	1						
X_{14}	0.940	0.101	0.746	0.919	1					
X_{23}	0.834	0.393	0.497	0.934	0.861	1				
X_{24}	0.923	0.116	0.714	0.859	0.931	0.932	1			
X_{34}	0.905	0.459	0.677	0.814	0.791	0.794	0.773	1		
X_{11}	0.788	0.210	0.516	0.931	0.929	0.848	0.844	0.673	1	
X_{22}	0.782	0.210	0.505	0.853	0.853	0.934	0.937	0.662	0.888	1

Table 12: Best model correlation coefficient matrix without insignificant variables

	Unstandardized coefficients			
Best model P26.5.3	B	Std. Error	t	Significance
(Constant)	-0.222	0.023	-9.776	4.752×10^{-17}
X_3	-4.090	0.284	-14.418	3.352×10^{-28}
X_{14}	0.186	0.011	16.601	3.389×10^{-33}
X_{23}	-3.452	1.321	-2.613	1×10^{-2}
X_{24}	0.400	0.026	15.297	3.067×10^{-30}
X_{34}	1.909	0.047	40.677	3.436×10^{-73}
X_{22}	-4.136	0.477	-8.667	2.135×10^{-14}

Table 13: Comparisons of the best pr models using newton's formula

Model	k+1	SSE	AIC	FPE	GCV	HQ	RICE	SCHWARZ	SGMASQ	SHIBATA
P26.5.3	7	0.059	5.05810^{-04}	5.05910^{-04}	5.07310^{-04}	5.38610^{-04}	5.09010^{-04}	5.90310^{-04}	4.80010^{-04}	5.03110^{-04}
P31.9.4	6	0.078	6.58210^{-04}	6.58310^{-04}	6.59710^{-04}	6.94610^{-04}	6.61210^{-04}	7.51410^{-04}	6.29210^{-04}	6.55610^{-04}
P32.10.3	7	0.063	5.37510^{-04}	5.37610^{-04}	5.39110^{-04}	5.72310^{-04}	5.40910^{-04}	6.27310^{-04}	5.10110^{-04}	5.34610^{-04}

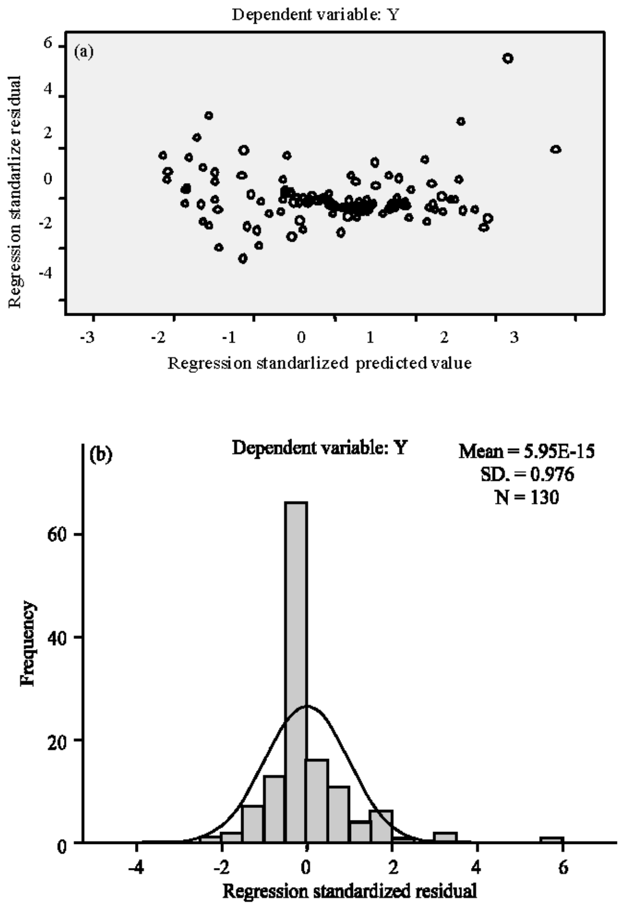

Fig. 6 (a-b): Scatter Plot and Histogram of the Regression Standardized Residuals of Model P26.5.3

The Goodness-of-Fit tests comprised of the randomness test and normality test. Randomness test was to determine that the residuals were normally distributed and normality test on the Kolmogorov-Smirnov statistics was to ensure that the normality assumptions were not violated. With a significance level of more than 0.05 (α>0.05), the normality test on the residuals gave the Kolmogorov-Smirnov statistics (0.192) of p-value >0.05 and the scatter plot of the standardized residuals showed that they were at random (Fig. 6). From the tests, the assumptions of randomness and normality of the residuals had therefore been satisfied.

The best polynomial regression model was thus given as in Table 12 by:

$$P26.5.3 = -0.22-4.09X_3-4.136X^2_2+0.186X_{14}-3.452X_{23}+0.400X_{24}-1.909X_{34} \quad (5)$$

Substituting the defined variables back into Eq. 5, the best model equation was thus:

$$P26.5.3 = -0.22-4.09D_D/T-4.136\,(D_m^{4.5})^2+0.186D_t^{3.7}T-3.452D_m^{4.5}D_b/T+0.400D_m^{4.5}T-1.909D_D \quad (6)$$

The model Eq. 5 showed the variable X_3 as main contributor and the significant contributions of the interaction terms (X_{14}, X_{23}, X_{24}, X_{34}) towards the stem volume. The major contributor to the stem volume was the diameter at the base (D_b) while the height was inversely proportional but interactively proportional with variables, diameter at the middle, D_m and the top, D_t. Diameters at the middle and the top had significantly contributed to the model as interaction growth factors while diameter at the base and stem height remained as main factors. The integers in the model equation ranged mathematically from 3.7-9.0, respectively. Hence, Eq. 5 and 6 signified the appropriateness of power transformation used in normalizing the variables before further regression analysis. Removal of multicollinearity and insignificant variables from the model had enhanced the model's efficiency based on the 8SC. The models goodness-of-fit had verified that the polynomial regression technique had thus improved on the stem volumetric estimation.

DISCUSSION

Multicollinearity effects in data sets would cause biasedness and inappropriate models being chosen for estimation and prediction. In this study, however, the problem of multicollinearity was addressed by initially normalizing and linearizing the data sets. The variables were then power transformed in the form of rational numbers ranging from values of 3.7-4.5. Since the best model had a quadratic term as in Eq. 5, thus the resultant model equation had a polynomial characterization of greater than 2 as shown in Eq. 6. Some previous researches had used only correlation analysis and coefficient of determination, R^2 in their studies (Nkongolo and Plassmeyer, 2010; FORMECU, 1997). High correlation coefficients between the independent variables exhibited perfect or strong collinearity. Without the removal of multicolhinearity source variables, this would cause biasedness in the models. Hence, this study had emphasized on these problems of multicollinearity and its removal from the models had able to enhance a model's performance based on their least residuals.

Previous studies had indicated the complexities of using polynomial regression of higher orders in a regression algorithm (Dam *et al*., 2000; Ekpenyong *et al*., 2008). However, in this study before any regression analyses were done, the data sets were transformed to normality. This would avoid an inappropriate model being chosen as the best model to be used for estimation and forecasting. Model's inappropriateness would later result in the violation of the regression assumptions of the residual analysis.

Remedial techniques in minimizing multicollinearity effects had been applied to obtain a robust model, followed further by the elimination of insignificant variables in the model. Multicollinearity removal had reduced the number of iterations needed to get to the best model and the elimination of the insignificant variables had further enhanced the models' performances. Statistical tests had been carried out to validate their removals. The Eight Selection Criteria (8SC) were effective in identifying the best model, where formally the criteria used in previous researches were based on the R^2 or the adjusted-R^2 for model selection (Daniel *et al*., 1979). Akindele and Le May (2006) had also used the standard error of estimate (SE) to predict the models overall performances. In this study, based on the SSE (0.059) of model P26.5.3, the eight selection criteria with each criterion of a penalty factor had able to validate the choice of model selection. Comparisons of the Newton's Polynomial Regression Models (PRM) based on the least 8SC, of which the errors ranged from 4.8×10^{-4}-5.903×10^{-4}, had appeared to represent an improved estimation for volumetric stem biomass estimation. Affendy *et al*. (2009) had indicated the good overall performance of *Cinnamomum iners* in survival rate (highest: 76.35) and growth increments (Diameter 0.98 cm year^{-1}; Height 0.77m year^{-1}) in the open planting technique. They also suggested that slashing, silvicultural treatment and continued monitoring could potentially improved the survival, height, diameter at breast height and biomass increment in a secondary forest area. In this study area where *Cinnamomum iners* were planted as shade trees as urban forest, landscaping recreational parks and along the boulevards and road-highways, tree stem dimensions, namely, diameters at the base (D_b), middle (D_m), top (D_t) and stem height (T) had implied significant contribution towards the stem volume estimation.

CONCLUSIONS

This study had tested the efficacy of polynomial regression to estimate the stem volume for forests stands. In this research, the polynomial relationships of the independent variables with the dependent could be attained easily by using the power transformation and the p-value approach of the normality tests on the variables. Model generation was done with diameters of trees as the determining growth. The best model P26.5.3 thus implied the significance of multicollinearity and insignificant variables removal in model enhancement for volumetric estimation. Diameter at the base and stem height were found to be the major contributors while other variables contribute significantly to the stem volume estimation. Stem volume estimation can therefore be carried out nondestructively, without the felling of trees, unless it was time for harvesting. The right time to coppice or harvest can also be predicted based on the mensurable variables. Other benefits can also be predicted such as, with proper sustainable forest management, revenues can be successfully generated based on the volumetric models. With selected species, for instance, commercially merchantable tree-logs like *Agarwood* or Aquilaria spp. that had been enlisted in the Endangered Species list in CITES, shortage and inconsistency of forest tree species in the world supply can therefore be overcome. Trees planted as urban or secondary forest can inadvertently, would improve the ecosystem and habitat as well as the gas cycles and carbon sequestration in the urban areas.

ACKNOWLEDGMENTS

The present research was prepared with the support from Universiti Malaysia Sabah as a part of Doctorate thesis. We also would like to thank Nigel Jonney J.B and Muhammad Faizal M.R. for their assistance in the field data collection.

APPENDIX

All possible polynomial models

P1 $V_1 = \beta_0+\beta_1X_1+\beta_{11}X_1^2+a_1$

P2 $V_2 = \beta_0+\beta_2X_2+\beta_{22}X_2^2+a_2$

P3 $V_3 = \beta_0+\beta_3X_3+\beta_{33}X_3^2+a_3$

P4 $V_4 = \beta_0+\beta_4X_4+\beta_{44}X_4^2+a_4$

P5 $V_5 = \beta_0+\beta_1X_1+\beta_2X_2+\beta_{11}X_1^2+\beta_{22}X_2^2+a_5$

P6 $V_6 = \beta_0+\beta_1X_1+\beta_3X_3+\beta_{11}X_1^2+\beta_{33}X_3^2+a_6$

P7 $V_7 = \beta_0+\beta_1X_1+\beta_4X_4+\beta_{11}X_1^2+\beta_{44}X_4^2+a_7$

P8 $V_8 = \beta_0+\beta_2X_2+\beta_3X_3+\beta_{22}X_2^2+\beta_{33}X_3^2+a_8$

P9 $V_9 = \beta_0+\beta_2X_2+\beta_4X_4+\beta_{22}X_2^2+\beta_{44}X_4^2+a_9$

P10 $V_{10} = \beta_0+\beta_3X_3+\beta_4X_4+\beta_{33}X_3^2+\beta_{44}X_4^2+a_{10}$

P11 $V_{11} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+a_{11}$

P12 $V_{12} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_4X_4+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{44}X_4^2+a_{12}$

P13 $V_{13} = \beta_0+\beta_1X_1+\beta_3X_3+\beta_4X_4+\beta_{11}X_1^2+\beta_{33}X_3^2+\beta_{44}X_4^2+a_{13}$

P14 $V_{14} = \beta_0+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+a_{14}$

P15 $V_{15} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+a_{15}$

P16 $V_{16} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{12}X_{12}+a_{16}$

P17 $V_{17} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_{11}X_1^2+\beta_{33}X_3^2+\beta_{13}X_{13}+a_{17}$

P18 $V_{18} = \beta_0+\beta_1X_1+\beta_4X_4+\beta_{11}X_1^2+\beta_{44}X_4^2+\beta_{14}X_{14}+a_{18}$

P19 $V_{19} = \beta_0+\beta_2X_2+\beta_3X_3+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{23}X_{23}+a_{19}$

P20 $V_{20} = \beta_0+\beta_2X_2+\beta_4X_4+\beta_{22}X_2^2+\beta_{44}X_4^2+\beta_{24}X_{24}+a_{20}$

P21 $V_{21} = \beta_0+\beta_3X_3+\beta_4X_4+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{34}X_{34}+a_{34}$

P22 $V_{22} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{12}X_{12}+\beta_{13}X_{13}+\beta_{23}X_{23}+\beta_{132}X_{132}+a_{22}$

P23 $V_{23} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_4X_4+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{44}X_4^2+\beta_{12}X_{12}+\beta_{14}X_{14}+\beta_{24}X_{24}+a_{23}$

P24 $V_{24} = \beta_0+\beta_1X_1+\beta_3X_3+\beta_4X_4+\beta_{11}X_1^2+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{13}X_{13}+\beta_{14}X_{14}+\beta_{34}X_{34}+a_{24}$

P25 $V_{25} = \beta_0+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{23}X_{23}+\beta_{24}X_{24}+\beta_{34}X_{34}+a_{25}$

P26 $V_{26} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{12}X_{12}+\beta_{13}X_{13}+\beta_{14}X_{14}+\beta_{23}X_{23}+\beta_{24}X_{24}+\beta_{34}X_{34}+\beta_{123}X_{123}+a_{26}$

P27 $V_{27} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_4X_4+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{12}X_{12}+\beta_{13}X_{13}+\beta_{23}X_{23}+\beta_{123}X_{123}+a_{27}$

P28 $V_{28} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{44}X_4^2+\beta_{12}X_{12}+\beta_{14}X_{14}+\beta_{24}X_{24}+\beta_{124}X_{124}+a_{28}$

P29 $V_{29} = \beta_0+\beta_1X_1+\beta_3X_3+\beta_4X_4+\beta_{11}X_1^2+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{13}X_{13}+\beta_{14}X_{14}+\beta_{34}X_{34}+\beta_{134}X_{134}+a_{29}$

P30 $V_{30} = \beta_0+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{23}X_{23}+\beta_{24}X_{24}+\beta_{34}X_{34}+\beta_{234}X_{234}+a_{30}$

P31 $V_{31} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{12}X_{12}+\beta_{13}X_{13}+\beta_{14}X_{14}+\beta_{23}X_{23}+\beta_{24}X_{24}+\beta_{34}X_{34}+\beta_{123}X_{123}+\beta_{124}X_{124}+\beta_{134}X_{134}+\beta_{234}X_{234}+a_{31}$

P32 $V_{32} = \beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_{11}X_1^2+\beta_{22}X_2^2+\beta_{33}X_3^2+\beta_{44}X_4^2+\beta_{12}X_{12}+\beta_{13}X_{13}+\beta_{14}X_{14}+\beta_{23}X_{23}+\beta_{24}X_{24}+\beta_{34}X_{34}+\beta_{123}X_{123}+\beta_{124}X_{124}+\beta_{134}X_{134}+\beta_{234}X_{234}+\beta_{1234}X_{1234}+a_{32}$

REFERENCES

Adekunle, V.A.J., 2007. Non-linear regression models for timber volume estimation in natural forest ecosystem, Southwest Nigeria. Res. J. For., 1: 40-54.

Affendy, H., M. Aminuddin. W. Razak, A. Arifin and A.R. Mojiol, 2009. Growth increments of indigenous species planted in secondary forest area. Res. J. For., 3: 23-28.

Akaike, H., 1970. Statistical predictor identification. Ann. Inst. Stat. Math., 22: 203-217.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Control, 19: 716-723.

Akindele, S.O. and V.M. Le May, 2006. Development of tree volume equations for common timber species in the tropical rainforest area of Nigeria. For. Ecol. Manage., 226: 41-48.

Alderete-Chavez, A., G.D.J. Zapata-Cocon, E. Ojeda-Trejo, E. Guevara and N.D.L. Cruz-Landero *et al*., 2010. Evaluation of growth *Cedrela odorata* L. in a secondary vegetation area. Res. J. For., 4: 145-150.

Ashish, K.S. and S. Muni, 1990. Regression Analysis: Thyeoriy, Methods and Applications. Springer, New York, USA.

Avery, T.E. and H.E. Burkhart, 1994. Forest Measurements. 4th Edn., McGraw-Hill, Inc., Singapore.

Bechtold, W.A., 2004. Largest-crown width prediction models for 53 species in the Western United States. West J. Applied For., 19: 245-251.

Cao, T., L. Valsta and A. Makela, 2010. A comparison of carbon assessment methos for optimizing timber production and carbon sequestration in Scots pine stands. For. Ecol. Manage., 260: 1726-1734.

Coakes, S.J. and L. Steed, 2007. SPSS: Analysis without Anguish: Version 14.0 for Windows. Wiley, Sydney, Australia.

Dam, J.S., T. Dalgaard, P.E. Fabrcius and S. Andersson-Engels, 2000. Multiple polynomial regression method for determination of biomedical optical properties from integratin sphere measurements. Applied Optics, 39: 1202-1209.

Daniel, V., J.A. Helms and F.S. Baker, 1979. Principles of Silviculture. 2nd Edn., McGraw-Hill Book Company, New York, pp: 500.

Devore, J. and R. Peck, 1993. Statistics: The Exploration and Analysis of Data. 6th Edn., Wadsworth, California, USA.,.

Ekpenyong, E.J., O.I. Mfon and J. Eno Donatus, 2008. Polynomial (Non Linear) regression method for improved estimation based on sampling. J. Applied Sci., 8: 1597-1599.

FORMECU, 1997. National Forest Resource Survey, Training Manual. Federal Department of Forestry, Abuja, Nigeria, pp: 47.

Fuwape, J.A., J.C. Onyekwelu and V.A.J. Adekunle, 2001. Biomass equations and estimation for gmelina arborea and nauclea diderrichii stands in Akure forest reserve. Biomass Bioenergy, 21: 401-405.

Golub, G.H., M. Heath and G. Wahba, 1979. Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics, 21: 215-223.

Gujarati, D.N., 2006. Essential of Econometrics. 3rd Edn., McGraw Hill Companies Inc., New York, USA..

Hannan, E.J. and B.G. Quinn, 1979. The determination of the order of an autoregression. J. R. Stat. Soc. Ser. B, 41: 190-195.

Hasenauer, H., 2006. Concepts within Tree Growth Modelling. Sustainable Forest Management. Springer, New York, USA.

Husch, B., T.W. Beers and J.A. Jr. Kershaw, 2003. Forest Mensuration. 4th Edn., John Wiley and Sons, Inc., New Jersey, USA., pp: 949.

John, D.S., 2003. Models for estimation and simulation of crown canopy cover. Proceedings of the 5th Annual Forest Inventory and Analysis Symposium. Nov. 18-20, New Orleans, LA. Gen. Tech. Rep. WO-69. Washington, D.C., U.S., pp: 222-222.

Kubo, T. and T. Kobayashi, 2007. Modelling of stem biomass accumulation in *Pinus densiflora* seedlings exposed to aqueous-phase OH radicals generating mist. Ecol. Modeling, 200: 467-474.

Nkongolo, N.V. and C.J. Plassmeyer, 2010. Effect of vegetation type on soil physical properties at lincoln university living laboratory. Res. J. For., 4: 1-13.

Noraini, A., H.J.J. Zainodin and J.B.N. Jonney, 2008. Multiple regression models of the volumetric stem biomass. WSEAS Trans. Math., 7: 492-502.

Peter, A.O. and O.A. Oluwafeni, 2008. Interim crown ratio models for a mixed tectona grandis and gmelina arborea stand in the university of Ibadan, Nigeria. Res. J. Forestry, 2: 34-42.

Philip, M.S., 1994. Measuring tress and Forests. 2nd Edn., CABI Publishing, United Kingdom Press, New York.

Ramanathan, R., 2002. Introductory Econometrics with Applications. 5th Edn., South-Western, Thomson Learning, Ohio, USA.

Rice, J., 1984. Bandwidth choice for nonparametric kernel regression. Ann. Statistics, 12: 1215-1230.

Schwarz, G.E., 1978. Estimating the dimension of a model. Ann. Stat., 6: 461-464.

Shawn, X.M., V.J. Lieffers and H. Shongming, 2007. Modelling crown volume of lodgepole pine based upon the uniform stress theory. For. Ecol. Manage., 251: 174-181.

Shibata, R., 1981. An optimal selection of regression variables. Biometrika, 68: 45-54.