An Improved Volumetric Estimation Using Polynomial Regression

Noraini Abdullah, Amran Ahmed and Zainodin Hj. Jubok

School of Science & Technology, Universiti Malaysia Sabah,
Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia

Corresponding emails: noraini@ums.edu.my, amran@ums.edu.my, zainodin@gmail.com

Abstract

The polynomial regression (PR) technique is used to estimate the parameters of the dependent variable having a polynomial relationship with the independent variable. Normality and non-linearity exhibit polynomial characterization of power terms greater than 2. Polynomial Regression models (PRM) with the auxiliary variables are considered up to their third order interactions. Preliminary, multicollinearity between the independent variables is minimized and statistical tests involving the Global, Correlation Coefficient, Wald, and Goodness-of-Fit tests, are carried out to select significant variables with their possible interactions. Comparisons between the polynomial regression models (PRM) are made using the eight selection criteria (8SC). The best regression model is identified based on the minimum value of the eight selection criteria (8SC). The use of an appropriate transformation will increase in the degree of a statistically valid polynomial, hence, providing a better estimation for the model.

Keywords: stem volume, polynomial regression models (PRM), normality, multicollinearity, ` eight selection criteria (8SC).

1. INTRODUCTION

According to Bechtold (2004), the stem (or bole) diameter was statistically significant and the strongest predictor in all models of tree estimations, and for some species a quadratic term was needed for model enhancement. While Hasenaeur (2006) had used modelling to predict future forest stand development, Noraini *et al.*(2008) had presented for tree stem biomass prediction and estimation using multiple regression models.

Stem Biomass Volumetric Equation

The variables considered during field data collection of 130 trees are diameter at the base (D_b), diameter at the middle (D_m), diameter at the top (D_t), and the stem height (T). The cylindrical area of log (A) is 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 then 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.

Area $=\pi R^2$ /bole where R is radius of stem

$=\pi D^2/4$ where D is the diameter

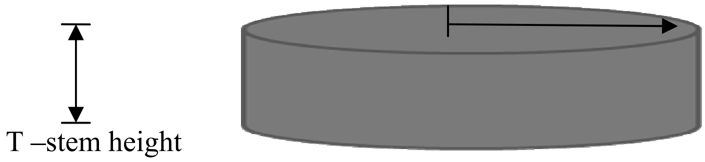


Figure 1. Merchantable Tree Log

The volume of the merchantable log will be the stem height multiplied by the area as shown in Figure 1 above. In this paper, based on these mensuration variables, the volume of the stem biomass is then calculated using the Newton's formula as in (1).

Newton's Formula (Fuwape *et al*., 2001):

$$V_{Nw} = \frac{T}{6}(A_b + 4A_m + A_t) \quad (1)$$

The objective of this study is to compare models based on the Newton's Formula using the polynomial regressions (PR) technique with significant attribution to the power terms.

2. METHODOLOGY

2.1 Data Preparation and P-value Method for Normality

In regression analysis, normality and linear relationship of data is of prime importance. Hence, normality test is initially carried out numerically (Coakes & Steed, 2007), with the graphical method as supporting evidence (Ashish & Muni, 1990). Normality tests are carried out in SPSS and the test statistic of the variable is given by the statistic value of Kolmogorov-Smirnov (for $n>50$) and Shapiro Wilk (for $n<50$). The confidence level is set at 95% with significance of α at 5% which is the standard percentage of the normality test.

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

Transformation will involve: i) identifying the types of curvilinear data, ii) determining the types of transformation needed, and iii) exercise the procedures for Ladder-Power or Box-Cox transformations. The ladder transformation procedure uses the data sets the power of the origin is employed, which is given by: (Devore & Peck, 1993)

$$\text{Transform value} = (\text{Original Value})^{\text{power}} \qquad (2)$$

Using the p-value from the F-statistics, data with p-value>0.05are considered as normal. Several iterations are executed so to determine the best transformation required for normality. Figure 2 depicts the flowchart on the data transformation procedures executed on non-normal or nonlinear data before any model building can be developed.

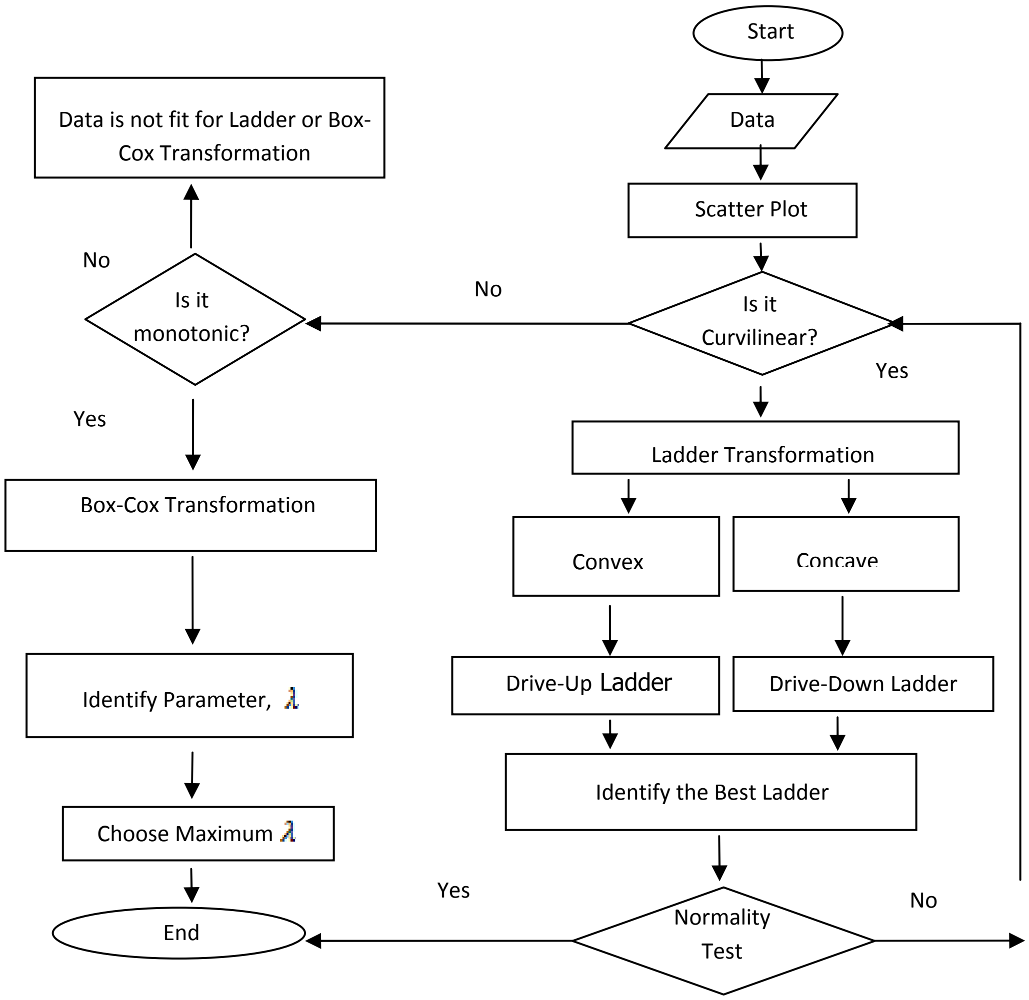


Figure 2. Flow Chart on the Procedures of Data Transformations

2.2 Modelling and Model-Building Approach

Figure 3 depicts the modelling flowchart. Preliminary with the conceptual development of the importance of modelling, its estimations, and contributions to the real world problems, mathematical theories are applied for model building.

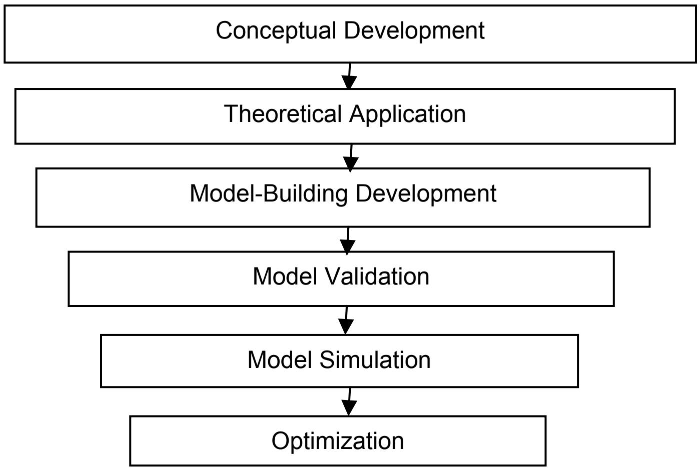


Figure 3: Modelling Flowchart

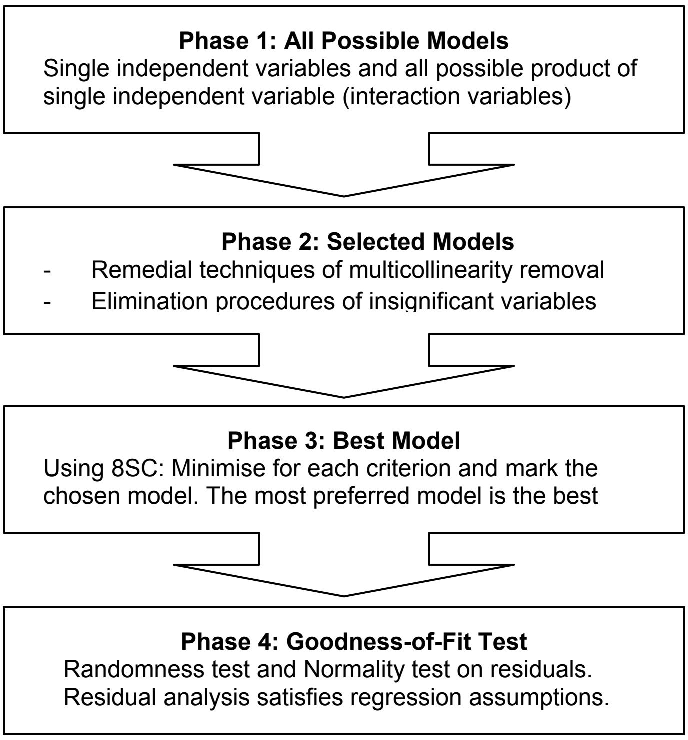


Figure 4. The Four Phases in Model-Building Development

Figure 4 shows the four phases of the Model-building development. Model-building techniques are exemplified and validated through tests and hypotheses. Model's validation is enhanced by simulation and optimization of values, expected to be characterized as optimal values. In this paper, the phases in model development will not be illustrated since the elimination procedures had been shown by Noraini *et al.,* (2008) and the multicollinearity removal techniques (Noraini *et al.*, 2010(a)).

2.3 Polynomial Regression Models (PRM)

Phase 1 of Model-building in Figure 4, consists of the all possible models which are made up of variables that have been prepared after undergoing the data preparation procedures of Figure 2. For simplicity, these variables are then known as the defined transformed variables.

The PR models are made of a dependent variable, V, the stem volume and single independent variables, taken from field data mensuration. The model-building is developed based on the method of multiple regressions, a statistical method of more than two independent variables as in (3),

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i \quad (3)$$

, where i=1, 2, …, n; Y_i is the dependent variable; X_{1i}, X_{2i}, …, X_{ki} are the independent variables; $\beta_i's$ are the regression coefficients with k parameters and ε_i are the residuals. As with polynomials of the order 2 (parabolic curve with quadratic terms), the model equation can be written as:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_{11} X_{1i}^2 + \ldots + \beta_k X_{ki} + \beta_{kk} X_{ki}^2 + \varepsilon_i \quad (4)$$

Based on say four single independent variables, the number of models then is 32 models (as shown in Table 1).

Table 1: Total Number of Possible Models

Number of Variables	Single Independent Variables	Order of Interactions			Total number of models
		1st	2nd	3rd	
1	4	-	-	-	4
2	6	6	-	-	12
3	4	4	4	-	12
4	1	1	1	1	4
Total	15	11	5	1	32

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

Table 2: All Possible Models of Four single independent Variables.

P1	$V_1 = \beta_0 + \beta_1 X_1 + \beta_{11} X_1^2 + \varepsilon_1$
P2	$V_2 = \beta_0 + \beta_2 X_2 + \beta_{22} X_2^2 + \varepsilon_2$
P3	$V_3 = \beta_0 + \beta_3 X_3 + \beta_{33} X_3^2 + \varepsilon_3$
P4	$V_4 = \beta_0 + \beta_4 X_4 + \beta_{44} X_4^2 + \varepsilon_4$
:	: : :
P15	$V_{15} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 +$ $\beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \varepsilon_{15}$
P16	$V_{16} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2$ $+ \beta_{12} X_{12} + \varepsilon_{16}$
:	: : : : : :

P27	$V_{27} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{23} X_{23} + \beta_{123} X_{123} + \varepsilon_{27}$
:	: : : : : :
P32	$V_{37} = \beta_0 + \beta_1 X_1 + \ldots + \beta_{11} X_1^2 + \ldots + \beta_{12} X_{12} + \ldots + \beta_{34} X_{34} + \beta_{123} X_{123} + \ldots + \beta_{1234} X_{1234} + \varepsilon_{32}$

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

$$V_{27} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{23} X_{23} + \beta_{123} X_{123} + \varepsilon_{27} \quad (5)$$

, 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} is the 2nd order interaction, and X_1^2, X_2^2 and X_3^2 as the polynomial term of power 2 (or also known as the quadratic terms).
The models can then be written in a general form as:

$$V_{PR} = \Omega_0 + \Omega_1 W_1 + \Omega_2 W_2 + \ldots + \Omega_k W_k + u \ , \quad (6)$$

, where V_{PR} is the volume, '*W*' is an independent variable which represents one of these types of variables, namely, single independent, interactive, generated, transformed, quadratic terms or even dummy variables, Ω's are the newly defined regression coefficients, and 'u' as the error terms for each respective transformed model. The number of models will depend on the number of single independent variables, given by the formula $\sum_{j=1}^{q} j({}^{q}C_{j})$ where '*q'* is the number of single independent variables.

2.4 Multicollinearity Removal and Insignificant Variable Elimination

Multicollinearity is a phenomenon where there exists very strong linear or perfect relationships between the independent variables (Gujarati, 2006), and collinearity between the variables can be identified by examining the values of the correlation matrix of the independent variables. High correlation coefficients of absolute values in the range of $0.75 \geq |r| \geq 0.95$ are considered to exhibit multicollinearity effects. These multicollinearity source variables have to be dealt with first before modelling can be done, as indicated in Phase 2 of model development in Figure 3. The elimination of insignificant variables from the models is carried out using the backward elimination method. Illustrations of the backward elimination method had also been shown by Noraini *et al.* (2008). In this paper, multicollinearity source variables with high correlation coefficient of absolute values greater than 0.95 ($|r| \geq 0.95$) are removed. The Case Types for multicollinearity removal procedures had also been illustrated by Noraini *et al.* (2010)(b).

2.5 Best Model Selection and Goodness-of-Fit Tests

Many criteria have been presented in order to select the best regression model, but none can be considered as the best one. Table 3 depicts the selection of the eight criteria (8SC) of Phase 3, used in identifying the best regression model (Ramanathan, 2002). The criteria are based on the value of sum of square error (SSE) where n is the number of samples or observations, and k+1) is the number of parameters in each respective model. The model having the least value in majority of the criteria will be chosen as the best model.

Table 3. Eight Selection Criteria (8SC) for Best Model Identification

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 & Quinn, 1979)	$\left(\frac{SSE}{n}\right)\ln n^{2(k+1)/n}$	SHIBATA (Shibata, 1981)	$\left(\frac{SSE}{n}\right)\frac{n+2(k+1)}{n}$

The best model will undergo the goodness-of-fit tests of Phase 4 in Figure 4, which comprises of the normality and randomness tests on the models' residuals. Without violating the assumptions in regression analysis, further simulations of the best model will provide a better prediction for future forest planning strategy and management.

3. MODELING RESULTS AND ANALYSES

3.1 Normality and Descriptive Statistics

The data variables are measured from 130 trees non-destructively, as defined in Table 3. Normality tests are done and transformations are carried out using Ladder-Power on the non-normal data. Table 4 depicted the defined variables, before and after transformations.

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
D_t	Diameter at top of trunk	$D_t^{3.7}$	X1
D_m	Diameter at middle of trunk	$D_m^{4.5}$	X2
D_b	Diameter at the base of trunk	D_b/T	X3
T	Tree height (m)	T	X_4

From Table 5, the p-values of variable D_t increase in the variable power range of 1.5–3.5, before decreases to the value of 4.5. The optimal (highest) p-value is 0.034, and the variable power is thus focused at 3.5.

Table 5. Normality Test Using Kolmogorov Smirnov on Variable D_t

Transformed Variable	Kolmogorov-Smirnov 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

Transformed Variable	Kolmogorov-Smirnov 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

Transformation power range is then chosen between 3.5- 4.5. Referring to Table 6, variable D_t has reached the optimal normality value of 0.061(highest) at the transformation value of 3.7. The second decimal digit will lie between 3.7-3.8. Similar procedures are executed on the other variable, D_m, and a generated variable, D_b/T_h, has been created for normality. Table 7 below depicts the descriptive statistics of the models' transformed variables. All the transformed variables have turned to normal since the significant p-value are more than 0.05. The data sets can then be used for further regression analysis.

Table 7. Descriptive Statistics of Transformed Variables

Defined Variables	Transformed Variables				
	V	X1	X2	X3	X4
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	0.18	0.01	0.00	0.07	3.78
Maximum	1.96	0.40	0.33	0.15	8.23
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.					

3.2 Multicollinearity Removal and Backward Elimination Method

In Phase 2 of model-building, multicollinearity source variables with high correlation coefficient of absolute values greater than 0.95 ($|r| \geq 0.95$) are thus removed. The Case Types for multicollinearity removal procedures had been illustrated by Noraini *et al.* (2010)(b).The elimination of insignificant variables from the models is carried out using the backward elimination method.

These procedures employed in Phase 2 will not be dealt with in detail, but then suffices 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, Table 9 and Table 10) respectively. The highlighted values in Table 8 indicate examples of high correlation values exhibiting multicollinearity effects of the independent variables (X1, X2, X12) which then result in the first multicollinearity removal of variable X12.

Table 8: Correlation Coefficient Matrix of Model P26.0

	Y	*X1*	*X2*	*X3*	*X4*	*X12*	*X13*	*X14*	*X23*	*X24*	*X34*	*X11*	*X22*	*X33*	*X44*
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

Subsequent five multicollinearity source variables (X1, X2, X12, X11,X44) are being removed have resulted in the correlation coefficient matrix of model P26.5.0 as shown in Table 9. Table 9 also shows the absence of high multicollinearity variables in the model where there are no more correlation coefficients of more than 0.95 exist in the model. The next step will be the process of eliminating insignificant variables from the model using the backward elimination method.

Table 9: Correlation Coefficient Matrix of Model P26.5.0

	Y	*X3*	*X4*	*X13*	*X14*	*X23*	*X24*	*X34*	*X11*	*X22*
Y	1									
X3	0.115	1								
X4	0.858	-0.331	1							
X13	0.840	0.412	0.502	1						
X14	0.940	0.101	0.746	0.919	1					
X23	0.834	0.393	0.497	0.934	0.861	1				
X24	0.923	0.116	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	0.788	0.210	0.516	0.931	0.929	0.848	0.844	0.673	1	
X22	0.782	0.210	0.505	0.853	0.853	0.934	0.937	0.662	0.888	1

The procedures of eliminating insignificant variables are then carried out as indicated in Table 10 below. Insignificant variables having the highest p-value or the least absolute value of the t-statistics will be eliminated. It can be seen that variables (X_{13}, X_4, and X_{11}) are subsequently to be removed since having p-values of more than 0.05, and hence they are not significant. Table 11 depicts the final matrix for the best model whereby all the remaining variables in the model are significant with their p-values less than 0.05 ($\alpha \leq 5\%$).

Table 10: Insignificant Variables Eliminated From Model P26.5.0

Models		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	t	Sig.	Action Taken
P26.5.1	(Constant)	-.113	.101	-1.120	.265	Highest p-value (0.795) or least absolute value of t-statistics |0.261|. Variable X13 is eliminated
	X3	-4.953	.891	-5.561	.000	
	X4	-.021	.018	-1.122	.264	
	X13	-.993	3.807	-.261	.795	
	X14	.254	.072	3.533	.001	
	X23	-2.707	4.385	-.617	.538	
	X24	.343	.083	4.158	.000	
	X34	2.067	.164	12.634	.000	
	X11	-.949	.751	-1.264	.209	
	X22	-3.126	1.033	-3.024	.003	
P26.5.2	(Constant)	-.112	.101	-1.115	.267	Highest p-value (0.267) or
	X3	-4.975	.883	-5.631	.000	
	X4	-.020	.018	-1.116	.267	
	X14	.239	.043	5.621	.000	

	X23	-3.786	1.453	-2.605	.010	least absolute value of t-statistics |1.116|. Variable X4 is then eliminated.
	X24	.360	.052	6.949	.000	
	X34	2.066	.163	12.681	.000	
	X11	-.960	.747	-1.285	.201	
	X22	-3.121	1.029	-3.032	.003	
P26.5.3	(Constant)	-.222	.023	-9.786	.000	Highest p-value (0.256) or least absolute value of t-statistics |1.140|. Variable X11 is eliminated
	X3	-4.042	.286	-14.115	.000	
	X14	.232	.042	5.511	.000	
	X23	-3.165	1.344	-2.355	.020	
	X24	.350	.051	6.853	.000	
	X34	1.893	.049	38.683	.000	
	X11	-.844	.740	-1.140	.256	
	X22	-3.095	1.030	-3.005	.003	

It can also be seen from Table 11 that only one single variable (X3), four first order interaction variables (X14, X23, X24, X34), and one variable of the polynomial (quadratic) term (X22).

Table 11: Correlation Coefficient Matrix of Best Model P26.5.3

Model P26.5.3	Unstandardized Coefficients		t	Sig.
	B	Std. Error		
(Constant)	-0.222	0.023	-9.776	4.752×10^{-17}
X3	-4.090	0.284	-14.418	3.352×10^{-28}
X14	0.186	0.011	16.601	3.389×10^{-33}
X23	-3.452	1.321	-2.613	1×10^{-2}
X24	0.400	0.026	15.297	3.067×10^{-30}
X34	1.909	0.047	40.677	3.436×10^{-73}
X22	-4.136	0.477	-8.667	2.135×10^{-14}

3.3 Best Model Regression Equation

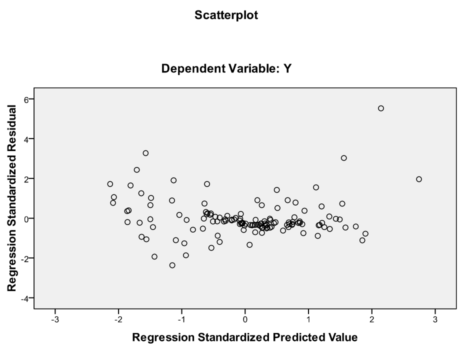
The best model from the 8SC is based on the (k+1) parameters, and fulfills the least value of most of the criteria (Ramanathan, 2002). Table 12 signifies the comparisons of the PR models based on the eight selection criteria. It can be seen the best PR model is represented by the model P26.5.3 with five multicollinearity removals and three insignificant variables eliminated.

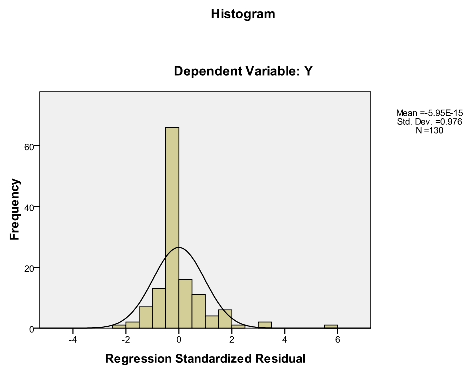
Table12. Comparisons of the Best PR Models Using Newton's Equation

Model	k+1	SSE	AIC	FPE	GCV	HQ	RICE	SCHWARZ	SGMASQ	SHIBATA
P26.5.3	7	0.059	5.058 E^{-04}	5.059 E^{-04}	5.073 E^{-04}	5.386 E^{-04}	5.090 E^{-04}	5.903 E^{-04}	4.800 E^{-04}	5.031 E^{-04}
P31.9.4	6	0.078	6.582 E^{-04}	6.583 E^{-04}	6.597 E^{-04}	6.946 E^{-04}	6.612 E^{-04}	7.514 E^{-04}	6.292 E^{-04}	6.556 E^{-04}
P32.10.3	7	0.063	5.375 E^{-04}	5.376 E^{-04}	5.391 E^{-04}	5.723 E^{-04}	5.409 E^{-04}	6.273 E^{-04}	5.101 E^{-04}	5.346 E^{-04}

The goodness-of-fit tests comprises of the randomness test and normality test. Randomness test is to determine that the residuals are normally distributed and normality test on the Kolmogorov-Smirnov statistics is to ensure that the normality assumptions are not violated. Since the sample size is 130, the random statistic, R is based on the normal (z) distribution. The null hypothesis is accepted since model P26.5.3 has zero mean of the residuals as shown by the scatterplot of the standardized residuals in Table 13. This implies that the residuals are independent and randomly distributed.

Table 13: Scatterplot and Histogram of the Regression Standardized Residuals.





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 (0.052) >0.05. From the good-of-fit tests and the plots, the assumptions of randomness and normality of the residuals have therefore been satisfied.

The best polynomial regression model is thus given by Table 11 as:-

$$P_{26.5.3} = -0.22 - 4.09X_3 - 4.136X_2^{\,2} + 0.186X_{14} - 3.452X_{23} + 0.400X_{24} - 1.909X_{34} \quad (7)$$

Substituting the defined variables back into equation (7), the best model equation is thus:

$$P_{26.5.3} = -0.22 - 4.09D_b/T - 4.136(D_m^{\,4.5})^2 + 0.186D_t^{\,3.7}T - 3.452D_m^{\,4.5}D_b/T + 0.400D_m^{\,4.5}T - 1.909D_b \quad (8)$$

Equation (8) signifies the appropriateness of the power transformation used in normalizing the variables before regression analysis. The range of integers in the model equation is mathematically from 3.7-9.0.

4. DISCUSSIONS AND CONCLUSIONS

Power Transformation in the form of integers is executed to normalize and linearize the data sets. The resultant model equation has polynomial characterization greater than 2. Previous studies had indicated that complexities of using polynomial regression in regression algorithm where higher orders of the polynomials are concerned (Dam *et al.*, 2000; Ekpenyong *et al.,* 2008). The polynomial relationships of the independent variables with the dependent can be transformed using the p-value method of the normality tests on the variables. Remedial techniques in minimizing multicollinearity effects are applied to obtain a robust model, further followed by the elimination of insignificant variables in the model. The eight selection criteria is effective in identifying the best model, where formally the criteria used is based on the R^2 or the adjusted-R^2 for model selection. Comparisons between the Newton's multiple regression models by Noraini *et al.*(2008) and Noraini *et al.*(2010(b)), based on the least 8SC, have appeared to represent an improved estimation using polynomial regression models (PRM) for volumetric stem biomass. Diameters at the base, middle, top and tree height have again signified as the main contributors towards the stem volume estimation.

Acknowledgements
The authors would like to acknowledge the financial support from Universiti Malaysia Sabah in this study.

REFERENCES

Akaike, H. (1970). "Statistical Predictor Identification". Annal Institute of Statistical Mathematics **22**, pp.203-217.
Akaike, H. (1974). "A New Look at model Statistical Identification". IEEE Trans Auto Control **19**, pp.716-723.
Bechtold, W.A. (2004). "Largest-crown width prediction models for 53 species in the western United States". West Journal Appl. Forest **19**(4), pp.245-251.
Ashish, K.S. & Muni, S. (1990). Regression Analysis:Thyeoriy, Methods and Applications. Springer, New York.
Coakes, S.J. & Steed, L. (2007). SPSS:Analysis without anguish:version 14.0 for windows. Wiley, Sydney.
Dam, J.S., Dalgaard, T., Fabrcius, P.E. & Andersson-Engels, S. (2000). Multiple Polynomial Regression Method for Determination of Biomedical Optical Properties from Integratin Sphere Measurements. Applied Optics **39**(7).
Devore, J. & Peck, R. (1993).Statistics, the Exploration and Analysis of Data. Wadsworth, California.
Ekpenyong, J.E., Okonnah, M.I. & John E,D. (2008). Polynomial (Non-Linear) Regression Method for Improved Estimation Based on Sampling. Journal of Applied Sciences **8**(8):1597-1599.
Fuwape, J.A., Onyekwelu, J.C. & Adekunle,V. A.J. (2001). "Biomass Equations and Estimation for Gmelina Arborea and Nauclea Diderrichii Stands in Akure Forest Reserve". *Biomass and Bioenergy* **21,** pp.401-405.
Golub, G.H., Heath, M. & Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics **21,** pp.215-223.
Gujarati, D.N. (2006). Essential of Econometrics. 3rd Edition. New York. McGraw Hill Companies, Inc.
Hannan, E.J. & Quinn, B.G. (1979). "The Determination of the Order of an Autoregression". Journal of Royal Statistics Society, Series **41(B)**, pp.190-195.
Hasenauer, H. (eds.). (2006). "Concepts within Tree Growth Modelling". *Sustainable Forest Management.* Springer. New York.
Noraini, A., Zainodin, H.J. & Nigel Jonney J.B. (2008). "Multiple Regression Models of the Volumetric Stem Biomass". *WSEAS Transactions on Mathematics* **7(7)** pp. 492-502.
Ramanathan, R. (2002). Introductory Econometrics with Applications. 5th Ed. South-Western, Thomson Learning, Ohio.
Rice, J. (1984). "Bandwidth Choice for Nonparametric Kernel Regression". Annals of Statistics **12**, pp.1215-1230.
Noraini, A., Amran Ahmed & Zainodin, H.J. (2010)(a). "Urban Forest Sustainability using MR Models". American Journal of Environmental Sciences (submitted for publication).
Noraini, A., Zainodin, H.J. & Amran Ahmed. (2010)(b). "Comparisons between Huber and Newton's Multiple Regression Models for Stem Biomass Estimation". Malaysian Journal of Mathematical Sciences (submitted for publication).
Schwarz, G. (1978). "Estimating the Dimension of a Model". Annals of Statistics **6**, pp.461-464.
Shibata, R. (1981)."An Optimal Selection of Regression Variables". Biometrika **68**(1), pp.45-54.

APPENDIX

All Possible Polynomial Models

P1	$V_1 = \beta_0 + \beta_1 X_1 + \beta_{11} X_1^2 + \varepsilon_1$
P2	$V_2 = \beta_0 + \beta_2 X_2 + \beta_{22} X_2^2 + \varepsilon_2$
P3	$V_3 = \beta_0 + \beta_3 X_3 + \beta_{33} X_3^2 + \varepsilon_3$
P4	$V_4 = \beta_0 + \beta_4 X_4 + \beta_{44} X_4^2 + \varepsilon_4$
P5	$V_5 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \varepsilon_5$
P6	$V_6 = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{33} X_3^2 + \varepsilon_6$
P7	$V_7 = \beta_0 + \beta_1 X_1 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{44} X_4^2 + \varepsilon_7$
P8	$V_8 = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \varepsilon_8$
P9	$V_9 = \beta_0 + \beta_2 X_2 + \beta_4 X_4 + \beta_{22} X_2^2 + \beta_{44} X_4^2 + \varepsilon_9$
P10	$V_{10} = \beta_0 + \beta_3 X_3 + \beta_4 X_4 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \varepsilon_{10}$
P11	$V_{11} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \varepsilon_{11}$
P12	$V_{12} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{44} X_4^2 + \varepsilon_{12}$
P13	$V_{13} = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \varepsilon_{13}$
P14	$V_{14} = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \varepsilon_{14}$
P15	$V_{15} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \varepsilon_{15}$
P16	$V_{16} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{12} X_{12} + \varepsilon_{16}$
P17	$V_{17} = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{33} X_3^2 + \beta_{13} X_{13} + \varepsilon_{17}$
P18	$V_{18} = \beta_0 + \beta_1 X_1 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{44} X_4^2 + \beta_{14} X_{14} + \varepsilon_{18}$
P19	$V_{19} = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{23} X_{23} + \varepsilon_{19}$
P20	$V_{20} = \beta_0 + \beta_2 X_2 + \beta_4 X_4 + \beta_{22} X_2^2 + \beta_{44} X_4^2 + \beta_{24} X_{24} + \varepsilon_{20}$
P21	$V_{21} = \beta_0 + \beta_3 X_3 + \beta_4 X_4 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{34} X_{34} + \varepsilon_{34}$
P22	$V_{22} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{23} X_{23} + \beta_{123} X_{123} + \varepsilon_{22}$
P23	$V_{23} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{44} X_4^2 + \beta_{12} X_{12} + \beta_{14} X_{14} + \beta_{24} X_{24} + \varepsilon_{23}$
P24	$V_{24} = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{13} X_{13} + \beta_{14} X_{14} + \beta_{34} X_{34} + \varepsilon_{24}$
P25	$V_{25} = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{23} X_{23} + \beta_{24} X_{24} + \beta_{34} X_{34} + \varepsilon_{25}$
P26	$V_{26} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{14} X_{14} + \beta_{23} X_{23} + \beta_{24} X_{24} + \beta_{34} X_{34} + \beta_{123} X_{123} + \varepsilon_{26}$
P27	$V_{27} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{23} X_{23} + \beta_{123} X_{123} + \varepsilon_{27}$
P28	$V_{27} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{44} X_4^2 + \beta_{12} X_{12} + \beta_{14} X_{14} + \beta_{24} X_{24} + \beta_{124} X_{124} + \varepsilon_{28}$

P29	$V_{27} = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{13} X_{13} + \beta_{14} X_{14} + \beta_{34} X_{34} + \beta_{134} X_{134} + \varepsilon_{29}$
P30	$V_{27} = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{23} X_{23} + \beta_{24} X_{24} + \beta_{34} X_{34} + \beta_{234} X_{234} + \varepsilon_{30}$
P31	$V_{27} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{14} X_{14} + \beta_{23} X_{23} + \beta_{24} X_{24} + \beta_{34} X_{34} + \beta_{123} X_{123} + \beta_{124} X_{124} + \beta_{134} X_{134} + \beta_{234} X_{234} + \varepsilon_{32}$
P32	$V_{37} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{14} X_{14} + \beta_{23} X_{23} + \beta_{24} X_{24} + \beta_{34} X_{34} + \beta_{123} X_{123} + .\beta_{124} X_{124} + \beta_{134} X_{134} + \beta_{234} X_{234} + \beta_{1234} X_{1234} + \varepsilon_{32}$