Predicting growth and yield models for *Eucalyptus* species in Aek Nauli, North Sumatera, Indonesia

Siti Latifah¹, Teodoro Reyes Villanueva², Myrna Gregorio Carandang², Nathaniel Cena Bantayan³, Leonardo M. Florece³

¹Forestry Program Study, Sumatera Utara University, Medan, Indonesia
²Forest Resources Management, University of the Philippines Los Banos, Los Banos, Philippines
³Environmental Science, University of the Philippines Los Banos, Los Banos, Philippines

Email address:
sitilatifah.164@yahoo.co.id (S. Latifah), teodoro_villanueva@yahoo.com (TR. Villanueva), mge405@yahoo.com (M. G. Carandang), ncbantayan@gmail.com (NC. Bantayan)

To cite this article:

Abstract: This study was conducted primarily to develop a yield prediction model for *Eucalyptus* spp plantations in Aek Nauli, North Sumatera, Indonesia as a contribution to sustain development and management of forest plantations. Data for growth and yield model were collected from the inventory and permanent sample plots (PSPs). The data in this study used 650 rhombic plots consisting of 106 PSPs and 544 inventory plots with several variations of plot size. Stands’ features referred diameter, height, merchantable volume, age, species, spacing, site index, basal area, and density of *Eucalyptus* species. Models using initial age, specifically, model 2 was found consistently to be the best model in most *Eucalyptus* plantations. Among the models using initial and projection age, model 4 was the better one. Model 2 using original ages looks better than model 4 because of it is being more reliable and its sigmoid growth curve. Nonetheless, significant differences were noticed between different models for predicting the merchantable volume of *Eucalyptus* spp. Plantations. Growth and yield models can be used to identify the best growing species of *Eucalyptus* spp. *E. hybrid* is recommended for plantation in this study area because it had the highest of merchantable volume.

Keywords: *Eucalyptus*, Growth, Yield, Merchantable Volume, Forest Plantation

1. Introduction

Indonesia is the world’s largest exporter of tropical timber, generating income of US$ 5 billion annually, and more than 48 million hectares (55 per cent of its remaining forests) are earmarked for logging. During 1990-2005, the country lost more than 28 million ha of forest indicating annual change in forest cover of -1,871,400 ha or annual deforestation rate of 2.0 per cent. Total forest loss since 1990 is around 24.1% (Mongabay, 2006).

One potential method for reducing tropical forest destruction is the creation of environmentally sustainable plantations that can supply wood and paper pulp product in order to reduce demand for wood from virgin forest (Hartomo, 2002).

The Government of Indonesia is developing timber estates to increase forest sector wood products. Wood industries are increasing rapidly; they need more than 40 million m³ of wood production every year. This cannot be supplied from the country’s natural forests alone and Indonesia, thus, need developing forest plantations. These programs are executed by private companies with forest concession areas (e.g. Industrial Timber Estate or HTI, Hutan Tanaman Industry) or State Forest Corporations (e.g. Perhutani and Inhutani) (Pramono and Agung, 1996).

In this case, Indonesian government policy will help rapid increase in Eucalyptus plantations for industrial wood use. *Eucalyptus* is favored for its rapid growth, indigenous origin and multipurpose use. Its major uses in Indonesia are for pulp (PT. TPL, 2008). Sumatera. However, in Indonesia, the estimation of yield and growth model for *Eucalyptus* spp has not been adequately investigated.
2. Material and Method

2.1. Study Site

The study was conducted in Aek Nauli sector of the Industrial Timber Estate (Hutan Tanaman Industri - HTI) namely, PT. Toba Pulp Lestari (PT. TPL), from December 2008 - March 2009. It is in the territorial jurisdiction of Porsea subdistrict, Simulungan regency, North Sumatera Province, Indonesia (Figure 1). The study area whose plots were used in the study belongs to climate A based on Schmidt and Ferguson (1951). It has the highest annual rainfall in October and the lowest in August. Based on its topography, the study area has a rainfall range of 85 – 434 mm/month, it has an elevation range from 350 to 1400 msl and it has top soil depth ranges from 10 to 44 cm. The study area covered an area 189,975 ha with Aek Nauli sector has area 21341.8 ha (Figure 1). Geographically, it is located between 02° 40'00" to 02° 50'00" north latitude and 98° 50'00" to 99° 10'00" east longitude.

Figure 1. Location of study site
height, density and weight. Also, they can be measured in value, which is the measure of economic and social interest. In this study, it is concentrated on physical measures of growth and yield namely, merchantable volume, because it forms the basis of value measures.

There have been many general forms of standard volume equations presented by many authors for a variety of conditions. There is no specific equation which is generally applicable and valid for all species. When tree volume is expressed as a function of both height and diameter, the relationship is called standard volume equation.

The regression equations used the logarithmic transformation models which were modified from the best model developed by some research such as that of Silva et al. (2007), Pereira et al. (2006), Yue et al. (2007), Villar (2005), Asvandi (2000) Bi et al. (1994), Bennet (1970) as cited by Clutter (1983), Sullivan and Clutter (1972). For the purpose, the following equations were generated. The models tested are shown in Table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>In V = b0 + b1 ln BA + b2 ln (Age) + b3 (SI) + b4 (SD) + b5 (SL) + b6 (DD) + b7 (TI) + b8 (R) + b9 (S) + b10 (E) + b11 (DA) + b12 (DI) + b13 (DD) + b14 (TI)</td>
</tr>
<tr>
<td>2</td>
<td>In V = b0 + b1 ln BA + b2 ln (Age) + b3 (SI) + b4 (SL) + b5 (DD) + b6 (TI) + b7 (R) + b8 (S) + b9 (E) + b10 (DA) + b11 (DI) + b12 (DD) + b13 (TI)</td>
</tr>
<tr>
<td>3</td>
<td>In V = b0 + b1 ln BA + b2 ln (Age) + b3 (SI) + b4 (SD) + b5 (SL) + b6 (DD) + b7 (TI) + b8 (R) + b9 (S) + b10 (E) + b11 (DA) + b12 (DI) + b13 (DD)</td>
</tr>
<tr>
<td>4</td>
<td>In V = b0 + b1 ln BA + b2 ln (Age) + b3 (SI) + b4 (SD) + b5 (SL) + b6 (DD) + b7 (TI) + b8 (R) + b9 (S) + b10 (E) + b11 (DA) + b12 (DI) + b13 (DD)</td>
</tr>
<tr>
<td>5</td>
<td>In V = b0 + b1 ln BA + b2 ln (Age) + b3 (SI) + b4 (SD) + b5 (SL) + b6 (DD) + b7 (TI) + b8 (R) + b9 (S) + b10 (E) + b11 (DA) + b12 (DI) + b13 (DD)</td>
</tr>
<tr>
<td>6</td>
<td>In V = b0 + b1 ln BA + b2 ln (Age) + b3 (SI) + b4 (SD) + b5 (SL) + b6 (DD) + b7 (TI) + b8 (R) + b9 (S) + b10 (E) + b11 (DA) + b12 (DI) + b13 (DD)</td>
</tr>
</tbody>
</table>

where:
- V = merchantable volume in cubic meter per hectare
- V1 = merchantable volume in initial ages in cubic meter per hectare
- V2 = merchantable volume in projection ages in cubic meter per hectare
- BA = basal area in square meters per hectare
- TH = tree height in meter
- Age = original age in year
- SI = site index in meter
- SD = stand density (tree/ha)
- SL = original spacing in square meters
- T = depth of top soil in centimeter
- R = rain fall monthly

SL = slope in percent (%)
E = elevation

DA, DB, DC and DD = dummy variable for group of soil
ln = natural logarithm or logarithm to the base 2.718281828
b0, b1, ..., bn = regression coefficient
1, 2, 3, 4 = model using original age
5, 6 = model using initial and projection ages

One approach to simplifying multiple regression equations is the stepwise regression. The procedure is probably the most widely used method for subset regression analysis (ANALYSTSOFT, 2008).

All equations were tested to determine which one is the best predictor of growth and yield for *Eucalyptus* species. The merchantable volume models were tested and respective fits of each the models were mainly assessed using criteria as follows:

1. The best model is the one which gives the largest coefficient of determination (R²) and coefficients of correlation (r), smallest mean square error (MSE) and values of Mallow’s C(p) and Variance Inflation Factor (VIF).
2. Significant F-values as computed in the analysis of variance (ANOVA) that test the overall regression given the intercept term.
3. Consider the significance of each the independent variables.
4. Randomness of the residual as shown in the graph of the residual error values versus fitted merchantable volume values.
5. Consider the parsimony of parameters. It is a subset of the independent variables which have to be small enough to manage cost and analysis facilitated but large enough for adequate description, control or prediction. That means, when all the leading equations appear to fit the data equally well, the simplest equation should be chosen as the best model.

The criteria that determine the best model is weighed with each criterion in the same scale as follows:
Score of weighted = (R²) + (r) - (MSE) - (Cp) + (Σ independent variables are significant / Σ independent variables in the model) + (Σ independent variables are not multicollinearity / Σ independent variables in the model)

The model selected should provide the highest accuracy as the models are verified against a separate set of yield data. The validation of the models developed in this study was done by using the chi-square test to determine the accuracy of the prediction models. This method enables one to determine whether the models follow the real world, and be examined using chi-square test of goodness of fit. This method enables one to determine whether the models follow the real world, and can be examined using chi-square test of goodness of fit. Specifically, the form of the chi-square is as follows:
\[ \chi^2_c = \frac{n}{(\alpha, n-1)} \sum (O_i - E_i)^2 \]

where:
- \( \chi^2_c \) = the chi-square value calculated
- \( \alpha \) = level of significance
- \( n \) = number of observations
- \( O_i \) = observed value
- \( E_i \) = expected (true) value

3. Results and Discussion

3.1. Selecting the Best Model

The regression equations were tested using several candidate regressions from non-linear models. When the relationship is non-linear, the model should be based on the nature of the curve representing given relationship. Two general types of equations are commonly used in regression analysis involving variables that exhibit non-linear relationship namely polynomial and exponential equations. These equations can be transformed into linear equations thru polynomial or logarithmic transformation.

Forest growth models in this study involving the logarithm of merchantable volume are the dependent variable. The use of the logarithm of yield as dependent variable is a convenient way to mathematically express the interaction of the independent variables in their effect on yield. Stand age, in most yield analysis, has been expressed as a reciprocal to allow for the ‘leveling off’ (asymptotic) effect of yield with increasing age. Site index is not often transformed prior to fitting, but sometimes logarithmic or reciprocal transformations are employed. In some models, height of dominant stand has been used in conjunction with age and the variable site index eliminated. The measure of stand density is commonly subjected to logarithmic transformation, particularly in models employing basal area, but the exact form in which density is included is quite variable, especially for plantation models which utilize number of trees per unit area as an independent variable. (Avery and Burkhardt, 1994)

The regression equations used two groups of independent variables. One group (model 1, 2 and 3) used original age, the basal area, height, site index, stand density, spacing, top soil, rain fall, slope, elevation and type of soil as independent variable. The other group used initial and projection age, volume, the basal area, height, site index, stand density, spacing, top soil, rain fall, slope, elevation and type of soil in initial age as independent variables.

Among the models using initial age, model 2 of the stepwise procedure was found consistently to be the best model because it was accepted in almost Eucalyptus species groups and also the model had the highest of the coefficients of determination (R²) and coefficients of correlation (r), and the lowest mean square error (MSE) and Mallows’s C(p), almost all independent variables are significant and also almost the independent variables is not multicollinearity.

Similarly, among the models using initial and using initial and projection age, model 4 was the best (Table 2)

3.2. Validation

Model validation is possibly the most important step in the model building sequence. It is also one of the most overlooked. A good representation of how the effectiveness a yield model as a predictive tool is its behavior in a real situation. Figure 2 and 3 show the comparison of graph between the actual data and predicted data for the best models for all species group. The angle between X-axis and a line on the graph is nearly 45°. Likewise is the angle between Y-axis and a line on the graph.

Table 2. Coefficients of determination (R²), coefficients of correlation (r), mean square error (MSE), and Mallows’s C(p). P-value and VIF among the best models

<table>
<thead>
<tr>
<th>Species Group</th>
<th>Model Adequate</th>
<th>Precision</th>
<th>Test of Hypothesis on βνs</th>
<th>Mult Collinear (MC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R² R</td>
<td>MSE C(p)</td>
<td>P-value</td>
<td>VIF</td>
</tr>
<tr>
<td>E. hybrid</td>
<td>2 0.919 0.857</td>
<td>0.7105 2.4271</td>
<td>4.6 significant 6.0 not mec</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 0.910 0.954</td>
<td>0.2532 3.1891</td>
<td>8.9 significant 9.0 not mec</td>
<td></td>
</tr>
<tr>
<td>E. grandis</td>
<td>1 0.981 0.991</td>
<td>0.0341 2.5881</td>
<td>4.5 significant 3.8 not mec</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 0.933 0.976</td>
<td>0.0341 3.2016</td>
<td>4.6 significant 6.0 not mec</td>
<td></td>
</tr>
<tr>
<td>E. pellita</td>
<td>2 0.995 0.997</td>
<td>0.0055 2.9787</td>
<td>4.4 significant 2.4 not mec</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 0.982 0.991</td>
<td>0.0152 3.8920</td>
<td>5.5 significant 5.5 not mec</td>
<td></td>
</tr>
<tr>
<td>E. urophylla</td>
<td>2 0.925 0.862</td>
<td>0.5840 0.0174</td>
<td>4.1 significant 4.1 not mec</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 0.990 0.995</td>
<td>0.0259 18.8819</td>
<td>6.6 significant 4.0 not mec</td>
<td></td>
</tr>
<tr>
<td>Mixed Eucalyptus</td>
<td>3 0.913 0.956</td>
<td>0.3977 3.4995</td>
<td>5.6 significant 6.0 not mec</td>
<td></td>
</tr>
<tr>
<td>All species</td>
<td>2 0.903 0.850</td>
<td>0.7763 2.2983</td>
<td>8.8 significant 7.8 not mec</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 0.830 0.911</td>
<td>0.3775 3.6742</td>
<td>6.6 significant 6.0 not mec</td>
<td></td>
</tr>
</tbody>
</table>

mc: multicollinearity
The capability of any prediction model to give accurate estimates is the real test for any such model. It is not necessary to establish that a model is "perfect", which is probably impossible because all models are abstractions of reality. But it is important to establish that the model is correctly implemented and free of errors. To test the accuracy of the yield model whether the model particularity follows the real world or not, two sets of observations were used in testing the goodness of fit of the yield model using chi-square ($\chi^2$) test.

Based on the results of computation, it can be concluded that there is no significant difference between the actual and the predicted value of merchantable volume. This was because $p$-value of the chi-square ($\chi^2$) of the model using original ages for E. hybrid, E. grandis, E. pellita, E. urophylla, E. mix and all species observations, respectively is 0.1601, 0.7587, 0.9952, 0.6191, 0.9999 and 0.0557 while $p$-value of the chi-square ($\chi^2$) of the model using initial and projection ages for E. hybrid, E. grandis, E. pellita, E. urophylla, and all species observations, respectively is 0.0505, 0.707, 0.3279, 0.2934 and 0.0538. The $p$-value of the chi-square ($\chi^2$) of both group models is greater than the corresponding 0.05 level of significance.

The residual plot is the best model shown which offers a good example of what a problem-free plot should look like (Figure 4). The residual plot is the best model shown which offers a good example of what a problem-free plot should look like. There are no curved trends in the plot, the average of the residuals is zero, and the points are equally represented about the x-axis. This residual represents the difference between the observed response merchantable volume and the value predicted by the regression line. If the regression model represents the data correctly, the residuals should be randomly distributed around the line of error = 0 with zero mean.
3.3. The Application of the Best Model to Get the Best Growing Species

The timber properties of grown Eucalyptus in the past have proven to have problems, but with the current increased interest in biomass for fiber and energy production, Eucalyptus species may now have a commercially viable end use. Various Eucalyptus species have been planted in Aek Naoli sector with varying degrees of success. As mentioned in the beginning of this chapter, in this study area, the species are grouped into E. hybrid (Ehyb), E. grandis (Egra), E. pellita E. urophylla (E. uro), E. Mix and E all. For E. hybrid

\[ V = \exp \left[ 0.12043 + 0.46844 \left( \ln \text{BA}*\ln \text{TH} \right) - 1.84144 \left( \text{Age}^{-1} \right) + 0.03701 \left( \text{TS} \right) + 0.00129 \left( \text{R} \right) + 0.14059 \left( \text{DC} \right) \right] \]

For E. grandis

\[ V = \exp \left[ 0.30190 + 1.38912 \ln \text{BA} - 0.40322 \ln \left( \text{Age}^{-1} \right) + 0.02682 \left( \text{SI} \right) - 0.000054 \left( \text{N} \right) \right] \]

For E. pellita

\[ V = \exp \left[ 1.65117 + 0.42590 \ln \text{BA}*\ln \text{TH} - 0.93169 \left( \text{Age}^{-1} \right) + 0.00023 \left( \text{E} \right) \right] \]

For E. urophylla

\[ V = \exp \left[ -3.22314 + 0.79323 \left( \ln \text{BA}*\ln \text{TH} \right) - 0.00351 \left( \text{R} \right) + 0.00094771 \left( \text{E} \right) \right] \]

For mixed Eucalyptus

\[ V = \exp \left[ 1.98198 + 0.81678 \left( \ln \text{BA}*\ln \text{Age} \right) - 2.16647 \left( \text{Age}^{-1} \right) + 0.07019 \text{SI}(\text{Age}^{-1}) - 0.90034 \ln \left( \text{N} \right) + 0.13791 \left( \text{TS} \right) \right] \]

For E. all species

\[ V = \exp \left[ -1.07511 + 0.50096 \left( \ln \text{BA}*\ln \text{TH} \right) - 1.39383 \left( \text{Age}^{-1} \right) - 0.51483 \left( \text{SI}^{-1} \right) + 0.11307 \left( \text{SP} \right) + 0.04181 \left( \text{TS} \right) + 0.00103 \left( \text{R} \right) - 0.22406 \left( \text{DA} \right) \right] \]

Pereira et al. (2006) reported that the prediction equation for Eucalyptus grandis in central Western region of Minas Gerais Brazil, including age, site index and basal area were fitted. These models fitted were also found to be good enough to predict growth and yield in E. grandis plantations in North Sumatera, Indonesia (Aswandi, 2000).

Table 3. Analysis models of variance (ANOVA) of the different yield models for E. hybrid, E. grandis, E. pellita, E. urophylla, mixed Eucalyptus and all species.

<table>
<thead>
<tr>
<th>SV</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
<th>F-Crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>1240720</td>
<td>5</td>
<td>248.14</td>
<td>42.782</td>
<td>2.64E-2</td>
<td>2.220</td>
</tr>
<tr>
<td>Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>8688804</td>
<td>148</td>
<td>5800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9929520</td>
<td>1503</td>
<td>5000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the result of the analysis of variance using the six models, it can be concluded that there is a significant difference between models among them to predict the merchantable volume of eucalyptus species. This is because the value of test statistic F is 42.782 is greater than the corresponding tabular F value of 2.22 at 0.05 level of significance (Table 3). Consequently, each Eucalyptus species has the specific model to predict the yield.

When the result of ANOVA indicates significant difference between models among them, we need to look which group of means are different or not. Based on the result of Duncan test, there is no significance difference between models for mixed Eucalyptus, E. pellita, E. urophylla and all species while there is no significance difference between models for E. pellita, E. urophylla, E. hybrid, E. grandis, E. hybrid and all eucalyptus species. It is because p-values are 0.078 and 0.124, respectively which are greater than 0.05 of level significance.

Figure 7. Predicted volume for different Eucalyptus species.
Based on the specific model for every Eucalyptus spp. to predict the yield. Figure 7 shows the graph for each Eucalyptus species. This figure indicates that E. hybrid observations has the highest (230 m³ ha⁻¹) of merchantable volume in year 5 and E. petlina has the lowest (105 m³ ha⁻¹) of merchantable volume in year 5.

E. hybrid has the highest (230 m³ ha⁻¹) of predicted volume in year 5 with final model V = exp [0.12043 - 0.4684 + 0.03791 (TS) + 0.14059 (DC)]. It means that there is an increase in volume by 0.46844 per unit increase in multiple basal area with top height holding other independent variables constant. There is a decrease in volume by 1.84144 per unit increase in reciprocal age holding other independent variables constant. There is an increase in volume by 0.03791 per unit increase in depth of top soil holding other independent variables constant. There is an increase in volume by 0.0129 per unit increase in rainfall holding another independent variables constant and last there is a decrease in volume by 0.14059 per unit increase in dummy variable for group of soil holding another independent variables constant.

The wrong choice of the species on a given site, wrong site selection for a given species and lack of tending operations were the principal factors for the low yield. It is assumed that the highest of merchantable volume is E. hybrid. the recommended plantation for this study area is E. hybrid. This species is suitable, as suggested by International development agencies, because the breeding new crops is important for ensuring sustainable forest management by developing new varieties that are higher-yielding, resistant to pests and diseases, drought-resistant or regionally adapted to different environments and growing conditions.

4. Conclusions

In relation to the objectives of the study, there are several conclusions reached as follows:

1. Based on six criteria used in this study, Model 2 using original ages can be applied to accurately growth and yield at study site. Each Eucalyptus species have a specific model to predict growth and yield for the study site.

2. There are eight independent variables affecting the yield of Eucalyptus namely: basal area, height, age, site index, spacing, depth of top soil, rainfall and group of site.

3. The quality of predictions obtained from any type of growth and yield model will depend on the availability of suitable data to calibrate and to initialize the model before each simulation.

4. With the highest value of averages merchantable volume at 7 year, Eucalyptus hybrid is the best growing species to be planted at the study site.

Acknowledgement

This research is part of PhD thesis of the first author funded by SEAMO-BIOTROP DIPA 2008. The authors extend deeply acknowledgment to the Ministry of Education and Culture – Republics Indonesia for scholarship and support to accomplish this paper. Sincerely appreciation is also extended to anonymous reviewer for correction and comments.

References


**ORIGINALITY REPORT**

<table>
<thead>
<tr>
<th>SIMILARITY INDEX</th>
<th>INTERNET SOURCES</th>
<th>PUBLICATIONS</th>
<th>STUDENT PAPERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>17%</td>
<td>10%</td>
<td>9%</td>
<td>2%</td>
</tr>
</tbody>
</table>

**PRIMARY SOURCES**


Publication

2. agris.fao.org

Internet Source

3. espace.library.uq.edu.au

Internet Source

4. fhut.usu.ac.id

Internet Source

5. www.symfor.org

Internet Source

Exclude quotes: On
Exclude bibliography: Off
Exclude matches: < 2%