Modelling Total Height of *Eucalyptus grandis* Hill ex Maiden

S.M.C.U.P. Subasinghe*

*Department of Forestry and Environmental Science, University of Sri Jayewardenepura, Nugegoda, Sri Lanka*

Date Received: 17-01-2014  Date Accepted: 07-09-2014

Abstract

*Eucalyptus grandis* Hill ex Maiden (rose gum) is an introduced species to Sri Lanka from Australia. At present it has drawn attention of both public and private sectors due to the use as fuel, railway sleepers and sawn timber. This study focused on prediction of total height of *E. grandis* with age, which is an essential requirement in plantation management.

Data were collected from 26 even-aged plantations, covering all favourable regions of Sri Lanka for the growth of *E. grandis*. At first theoretical equations were formulated with four possible transformations. Then parameters were estimated by fitting data at three different stages. $R^2$ values and standard residual distributions were used as preliminary evaluations. Initially it was tried to model height using tree age as the only explanatory variable. Both linear and exponential functions were used at this stage. The resultant models, however, were not successful for both functions due to low $R^2$ values. Therefore next attempt was made after partitioning the data into different site types using a simple site index. Three different site types were identified at this stage. Then linear and exponential functions were separately fitted to each site type to estimate parameters for different site types while keeping the same basic equation forms. This attempt was also not successful due to low $R^2$ values, large outliers (for some site types) or incompatibility of the resultant models with biological reality.

After the above unsuccessful attempts, the last step was conducted by pooling the data again and incorporating a second explanatory variable, site index. Other than multiple linear functions, exponential and logistic functions were modified by adding the second explanatory variable at this stage. Based on $R^2$ and standard residuals, seven suitable models were selected. Then the estimated height values were fitted against an age series to test the distribution and compatibility with biological reality. Finally, after both qualitative and quantitative evaluations, the best model was selected to predict the height growth for *E. grandis* in Sri Lanka for all site types.

*Keywords: height model, Eucalyptus grandis, site index, mathematical modelling*

1. Introduction

Modelling is especially important for species of widespread commercial use, both to understand growth and development of the species and to make better management decisions aimed at increasing productivity (Fernandez and Norero, 2006). Moreover, accurate growth and yield predictions of trees and forests are important requirements for facilitating sustainable management of forest resources.

*Correspondence: upuls@sjp.ac.lk
Tel: +94 112804685
ISSN 2235-9370 Print/ISSN 2235-9362 Online © University of Sri Jayewardenepura*
In commercial forestry, the important management decisions on different activities such as fertilising, thinning and harvesting are taken long before the trees achieve the required end dimensions. Plantation age is therefore commonly considered as the first input parameter to predict the future values of important tree growth variables such as diameter, height, volume etc. It is therefore important to be able to make predictions of required tree variables from age, so that the change of growth with time can be readily determined.

A second important input variable for yield prediction can be the quality of the site where the forests are grown. Different sites support the tree growth in different manner, even for the same species. Therefore it is common to include simple tree variables to represent the site quality (e.g., Pienaar and Harrison, 1989; Soares et al., 1995) or site indices (Clutter et al., 1992; Vanclay, 1994), in the forest growth and yield models although the site classification can be done using many different methods (Clutter et al., 1992). Those simple variables or indices are frequently used in order to ease the measurement procedures and thereby to make the use of the constructed model is more practical. Dominant height or top height are the most frequently used tree variables in this respect as these are considered to be independent of inter-tree competition (Pienaar and Harrison, 1989; Clutter et al., 1992; Soares et al., 1995).

Prediction of dbh and height is important in forestry since these two variables represent the horizontal and vertical growth of the stem respectively. Height growth is an essential feature of most growth and yield models, which are a principal tool of forest management planning (Boisvenue et al., 2004). The time required for a tree to reach a given height varies with its current growth rate, species, site quality, geographic location, site attribute and competition (Carmean, 1975; Oliver and Larson, 1996). However, including all these factors in a growth model can be difficult due to difficulty of quantifying some of the parameters (e.g. competition) and due to the high cost of the required measurements.

Peterson and Peterson (1994) found that most of the variation in height growth in harsh environments was due to species differentiation, with site and aspect as contributing factors. Individual tree height growth partly changes the stand structure and in turn, stand structure is a determinant of individual tree height growth. Modelling of height can therefore be used to reveal underlying factors.

In forest growth modelling, it is common to find both construction of new mathematical equations (e.g. Fontes et al., 2003; Boisvenue et al., 2004; Wang et al., 2005) as growth models or development of existing mathematical equations further to achieve more realistic predictions. Among the already available mathematical functions which have been used to develop growth or yield models in the past, Lundqvist-Korf (1939), Schumacher (1939), Bertalanffy (1957) and Richards (1959) functions are most common (e.g. Palahi et al., 2004; Sanchez-Gonzales et al., 2005; Salas and Garcia, 2006; Rammig et al., 2007; Adame et al., 2008). The new models were mostly constructed by using assumptions on the relationships between the response variable and the explanatory variables. In latter stages of model building those relationships were mathematically tested to obtain the statistical parameters which determine the magnitude and the direction of the relationships.

In 1994, Niklas stated that simple linear relationships were appropriate for tree growth prediction. However, according to Fernadez and Norero (2006), more complex formulae can be utilised to describe morphological association of variables. This change in approach may result from the
increased availability of affordable computing and the development of easier to use statistical packages. However, it is important that the selected candidate variables for modelling should represent the basic biological processes of tree growth (Boisvenue et al., 2004).

2. Methodology
2.1 Details of the plantations selected

Rose gum (Eucalyptus grandis Hill ex Maiden) was selected for the current study since it plays a major role in plantation forestry in Sri Lanka. It is mostly managed as even-aged monocultures in the country in the areas of upcountry wet zone. In order to represent the areas where the selected species is mainly grown, 26 plantations were selected from Badulla, Bogawantalawa, Haputale, Kandapola, Kandy, Maskeliya, Nuwara Eliya and Pattipola regions. The age of those plantations varied from 9 to 45 years.

2.2 Sampling and data collection

Although even-aged monoculture plantations were selected for the present study, stratified random sampling was employed for data collection in order to have a better representation for the entire plantation. Stratification was visually done by observing geographical variations at each site. At least two samples of 0.02 ha were selected in each stratum.

2.3 Construction of model structures

Since height generally increases with age, it was decided to use age as the primary explanatory variable. Although it was initially assumed that the growth of height can be predicted primarily as a function of age, it was accepted that site quality and degree of inter-tree competition can both also have a major effect on tree growth rates. With considering the above factors, the constructed basic model structure is given in equation 1.

\[
\text{Height} = f (\text{age, site quality, competition})
\]

Similar model structures were used by Subasinghe (1998) and Lee et al., (2004) for constructing mathematical models in forestry. Lee et al., (2004), however, tried to eliminate the age from the selected explanatory variables, but found out that modelling performance decreased if fitted without age.

At the first stage, and in order to construct a simple model, it was decided to model height separately using age as the only explanatory variable. The use of other explanatory variables especially site quality was considered only if this exercise failed. Competition was eliminated as an explanatory variable since the tree’s own growth parameters already reflected the competition experienced.

2.4 Identification of the relationships between selected variables

According to many growth modellers (Soares et al., 1995; Vanclay, 2004; Wang et al., 2005), the best way to reveal the relationships between explanatory and response variables is to study the scatter distribution. Following their work, it was decided to observe the scatter distributions between selected variables in order to identify the pattern of relationships (i.e., whether linear or non-linear) of height with age and thereby to identify the sign associated with statistical parameters.
2.5 Data partition due to growth differences

In order to partition data, the scatter distribution of height with age was both carefully examined. Moreover, the distribution of top height with age was tested since top height is believed to be a good indicator of site quality since it is independent from the competition (Clutter et al., 1992; Philip, 1994). In order to combine the above qualitative results with a quantitative value, a site index function was developed using top height and age (equation 2).

\[
SI = \frac{h_{top}}{a}
\]

where: \(a\) = plantation age, yr
\(h_{top}\) = top height, m
\(SI\) = site index

2.6 Stages of model construction

The entire modelling exercise conducted in this study can be divided into three stages. In order to build a less complex model, it was decided to construct a simple model to predict height using age as the only explanatory variable at the first stage. Moreover, it was decided to construct a common model without partitioning the data by site or growth differences. The reason for this latter approach was to eliminate the complexity of using different models to predict the same variable for different site types.

If the models constructed at the first stage indicated bias, the next stage of the present study was to model height separately for the partitioned data by growth differences. After partitioning data according to different growth patterns, the next step was to fit the same model structure (linear or non-linear) separately to different data sets to obtain different sets of statistical parameters.

In order to simplify the constructed models, it was decided to first use simple linear models and then to compare these with exponential and logistic functions. If the attempts on both stage one and two were failed, the next possible option was to include an additional explanatory variable to represent the quality of site without partitioning data. For this reason, the age was selected as the first (primary) explanatory variable and in addition to that, a second variable, i.e., site quality, was decided to use.

Second explanatory variable (site quality)

The second variable, site quality, was used in two ways; as quantitative and partially qualitative variables. As a quantitative value for the site index, it was decided to use top height/plantation age (equation 2). When the mean top height/age for each plantation was observed, it was possible to identify three distinctive sets of values and thereby three different site classes. Then a new site index was developed accordingly for each class (see Table 1) by assigning a qualitative value (column 4 of Table 1) which is approximately similar to the average top height/age (column 3) value for each class. This value was deliberately reduced to 0.5 for the Class III in order to keep a constant interval between the new SI values. For the other two classes, the new SI values were similar to the calculated average values in the column 3. Finally each class is given a description as Good, Average or Poor classes.

Table 1: Newly assigned site classes for different site types.

<table>
<thead>
<tr>
<th>Class</th>
<th>Growth rate</th>
<th>Average SI</th>
<th>New SI</th>
<th>Description</th>
<th>Plantations</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>High</td>
<td>2.4</td>
<td>2.5</td>
<td>Good</td>
<td>4</td>
</tr>
<tr>
<td>II</td>
<td>Average</td>
<td>1.6</td>
<td>1.5</td>
<td>Average</td>
<td>18</td>
</tr>
<tr>
<td>III</td>
<td>Poor</td>
<td>1.1</td>
<td>0.5</td>
<td>Poor</td>
<td>4</td>
</tr>
</tbody>
</table>
It was decided to add the second explanatory variable (SI) to the model in two ways, *i.e.*, as an additive (equation 3) and multiplicative variable (equation 4). For each situation described in equations 3 and 4, the combination of the SI and associated statistical parameter was included in three ways, *i.e.*, (i) parameter \( \times \) SI; (ii) parameter \(^\text{SI}\); and (iii) SI

\[ h = f \text{age (as explanatory or logistic)} + (\text{top height/age or SI}) \]  
\[ h = f \text{age (as explanatory or logistic)} \times (\text{top height/age or SI}) \]

Age was included in exponential or logistic form into the models. It was therefore not transformed into any other form. However, SI (both top height/age and new SI) was transformed into four biologically acceptable forms, *i.e.*, natural logarithmic, square, square root and reciprocal.

2.7 Fitting models structures to data

Fitting the selected non-linear models to data was done using SPSS statistical software. The major forms of the models are given in the equations 5 and 6 (as explanatory models) and 07 and 08 (as logistic models) without transformation of the SI variable.

\[ h = b_1 \times b_2^{\text{age}} + b_3 \text{SI} \]  
\[ h = b_1 \times b_2^{\text{age}} \times b_3 \text{SI} \]

\[ h = c_1 \times (1 + \exp(-c_2 \times (\text{age} - c_3))) + c_4 \text{SI} \]  
\[ h = c_1 \times (1 + \exp(-c_2 \times (\text{age} - c_3))) \times c_4 \text{SI} \]

2.8 Model evaluation

The quality of the resultant modes were primarily examined using \( R^2 \) and distribution of residuals. For the selected models by above two tests, three quantitative measurements, *i.e.*, average model bias, mean absolute difference and modelling efficiency were calculated to evaluate the chosen models further. The latter tests have proved adequate for model evaluation by many authors in the past (Fontes et al., 2003; Hein et al., 2007; Rammig et al., 2007; Rodriguez et al., 2003; Soares et al., 1995).

2.9 Validation of the finally selected models

This step was done to identify the most suitable models to be used in the field. The dbh and height values were predicted using ages from year 5 to 50 at five-year intervals separately for three classes. For the models which contain top height/age as the second explanatory variable, the average values calculated at the model construction phase were used for each class (*i.e.*, 2.4, 1.6 and 1.1 – Table 1). For the models which contain SI values 2.5, 1.5, 0.5 were respectively used for Class I, II and III (Table 1) as the second explanatory variable.

3. Results

3.1 Stage one: construction of simple common models using age as the only explanatory variable

The simple linear model (equation 9) resultant at this stage was given in equation 9. \( R^2 \) of that model was 50.0% and the intercept was significant which could not biologically be explained without specifying a range of validity using age.
\[ h = 17.7 + 0.665 \times \text{age} \]

Transforming variables into other forms did not improve the quality of \( R^2 \) or residual distribution. Due to these reasons, the above model was therefore discarded. In order to be compatible with growth rates slowing with age, it was decided to use exponential relationships to predict height from age under the stage one. When the exponential function was fitted, the resultant \( R^2 \) value was low due to the high leverage of some data points. It was therefore decided to observe the improvement of the models after removing the datum which had the highest leverage. Even the new model (equation 10), had low \( R^2 \) value (56.1%). The fitted line plot for the equation 9 is given in the Figure 1.

\[ h = 51.8 - 47.1 \times (0.9548^{\text{age}}) \]

![Figure 1: Fitted exponential models with observed data for height after removing datum with the highest leverage.](image)

In addition to that, there was an over-estimation by the resultant model (equation 10) in the early years of age according to the fitted model (Figure 1). The logistic model built behaved in the same manner and therefore the stage one was not successful for the present study.

3.2 Stage two: modelling with partitioned data

In order to partition the data by the growth differences as described in the methodology, scatter distribution between height and age was observed (Figure 2). According to that figure, the distribution of height with age indicated three separate clusters. The majority of data were located at the middle and four plantations were located above (KP 1,43; PP 1,85; PO 4,11 and PO 2,16) and four plantations were located below (BK 2,10; BK 2,24; PP 1,08 and HT 2,08) the major cluster. These variations were identical for both dbh and height and therefore it was concluded that the growth of rose gum in measured plantations was being influenced by some other parameters in addition to the age. At this stage that factor was assumed as the site quality which was therefore used as the second explanatory variable in latter stages of model building.
According to those results, the entire data set was divided into three classes as shown in the Table 2. The average site index values calculated for different growth classes are also given in the same table. After partitioning data, the next step was to fit the same model structure (linear or non-linear) separately to the three sets of data to obtain the relevant statistical parameters.

Table 2: New categorisation of measured rose gum plantations according to partitioned data.

<table>
<thead>
<tr>
<th>Class</th>
<th>Growth rate</th>
<th>Average SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>High</td>
<td>2.4</td>
</tr>
<tr>
<td>II</td>
<td>Average</td>
<td>1.6</td>
</tr>
<tr>
<td>III</td>
<td>Poor</td>
<td>1.1</td>
</tr>
</tbody>
</table>

**Simple linear models**

The simple linear models constructed to predict the dbh and height using age and the associated $R^2$ values are given in Table 3.

Table 3: Resultant simple linear models for different site classes to predict height.

<table>
<thead>
<tr>
<th>Class</th>
<th>Model</th>
<th>$R^2$%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$19.370 + 1.0340 \times age$</td>
<td>93.9</td>
</tr>
<tr>
<td></td>
<td>$2.005 \times age^{15}$</td>
<td><strong>93.6</strong></td>
</tr>
<tr>
<td>II</td>
<td>$6.310 + 1.1043 \times age$</td>
<td>90.2</td>
</tr>
<tr>
<td></td>
<td>$1.348 \times age^{15}$</td>
<td>85.3</td>
</tr>
<tr>
<td>III</td>
<td>$0.961 \times age^{25}$</td>
<td>78.9</td>
</tr>
</tbody>
</table>

$^{15}$Intercept was significant, but forced through 0;  
$^{25}$Intercept was not significant, but residual distribution became poor when re-fitted without intercept.
The intercept of the height prediction models were significant for the models built for Classes I and II. Class III was the only occasion where the intercept was not statistically significant. When forcing the intercept to zero for the other two classes, i.e., I and II, the $R^2$ value dropped significantly especially for the Class I (Table 3) or the residual distribution became poor. Therefore the linear models built for different growth classes were not successful for this study.

**Prediction of height using age as an exponential function**

In order to eliminate the drawbacks of the linear models, it was decided to fit standard curves to predict height of rose gum plantations. As the first step, age was used as an exponential function without the asymptote. Since Class II has a larger data set, those data were modelled first. The resultant model is given in equation 11.

$$h = 14.180 \times 1.032^{age}$$

The estimated $R^2$ value for height model was 88.0% and the standard error was 4.12 m. The parameter associated with age was higher than 1.0 which showed an indefinite increase of height with age. This phenomenon cannot biologically be explained and therefore it was not decided to test the exponential equations for the next two classes.

**Prediction height using age as a logistic function**

In order to construct better models to predict height, age was included as a logistic function without the asymptote. The resultant models and $R^2$ values are given in Table 4. Other than the height prediction model for Class III, the resultant $R^2$ values were approximately 90% or greater (Table 4). However, the standard residual distribution for the models built for the growth class III, it was poor.

<table>
<thead>
<tr>
<th>Class</th>
<th>Model</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$43.93/(1+e(-0.193 \times (age-6.03)))$</td>
<td>99.1</td>
</tr>
<tr>
<td>II</td>
<td>$63.60/(1+e(-0.076 \times (age-23.93)))$</td>
<td>89.8</td>
</tr>
<tr>
<td>III</td>
<td>$50.50/(1+e(-0.142 \times (age-30.81)))$</td>
<td>56.3</td>
</tr>
</tbody>
</table>

Due to those reasons, the attempts made to build models using linear, exponential and logistic functions separately for different growth glasses to predict height was not successful at the stage one.

**3.3 Stage three: prediction of height using non-standard non-linear models**

Due to the failures of stage one and two, non-standard non-linear models were selected in order to improve the model predictions by adding more than one explanatory variable as described in the methodology. After carefully studying the residual distributions and the results of quantitative evaluation, it was possible to select 7 models from height prediction models as given in Table 5. Priority was given for the high modelling efficiency values and the low mean absolute differences since these are quantitative values. Only then were the normal residual distributions considered.
Table 5: Selected height prediction models for the validation.

<table>
<thead>
<tr>
<th>No</th>
<th>Model</th>
<th>Modelling Efficiency</th>
<th>Mean absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(\frac{43.852(1+\exp(-0.117\times(A-25.497)))+7.590^\times(SI^0.5))}{(1+\exp(-0.117\times(A-25.497)))+7.590^\times(SI^0.5))})</td>
<td>0.93</td>
<td>2.64</td>
</tr>
<tr>
<td>2</td>
<td>(\frac{41.832(1+\exp(-0.128\times(A-28.664)))+(11.098\times SI^3)}{(1+\exp(-0.128\times(A-28.664)))+(11.098\times SI^3)})</td>
<td>0.92</td>
<td>2.63</td>
</tr>
<tr>
<td>3</td>
<td>(\frac{47.474(1+\exp(-0.100\times(A-20.930)))+(3.165^\times SI)}{(1+\exp(-0.100\times(A-20.930)))+(3.165^\times SI)})</td>
<td>0.90</td>
<td>2.97</td>
</tr>
<tr>
<td>4</td>
<td>(\frac{32.005(1+\exp(-0.288\times(A-18.773)))+(8.221\times TopA)}{(1+\exp(-0.288\times(A-18.773)))+(8.221\times TopA)})</td>
<td>0.90</td>
<td>2.62</td>
</tr>
<tr>
<td>5</td>
<td>(\frac{46.772(1+\exp(-0.114\times(A-17.085))\times(0.908\times(TopA^0.5))}{(1+\exp(-0.114\times(A-17.085))\times(0.908\times(TopA^0.5))})</td>
<td>0.90</td>
<td>2.78</td>
</tr>
<tr>
<td>6</td>
<td>(\frac{11.240\times(1.032^\times A)\times(1.112^\times(SI^2))}{(1.032^\times A)\times(1.112^\times(SI^2))})</td>
<td>0.91</td>
<td>2.62</td>
</tr>
<tr>
<td>7</td>
<td>(\frac{72.402(1+\exp(-0.051\times(A-36.728))\times(1.107^\times(SI^2))}{(1+\exp(-0.051\times(A-36.728))\times(1.107^\times(SI^2))})</td>
<td>0.89</td>
<td>2.92</td>
</tr>
</tbody>
</table>

Validation of the selected models

Distribution of predicted values of height of the models listed in Table 5 with a series of age is given in Figure 3.

3.4 Final evaluation of the height prediction models

When observing Figure 3a, b and c, it was clear that the models h6 and h7 over-estimated the height values with age. Both models were therefore removed from further study. Model h4 indicated the complete stoppage of height growth after 25 years of age which is highly unlikely in Sri Lankan conditions. The model h4 was therefore eliminated from further study. When the residual distributions were examined, those were poor for the models h1 and h3. Therefore those models were also eliminated.

3.5 The chosen models

Both the selected models (h5 and h2 in Table 5) contained age expressed logistically and the difference is given by the type of the second explanatory variable and the way it was used in the models. These models are re-written as equations 12 and 13 (h5 and h2 respectively in the Table 5) below and assigned the new numbers as models 1 and 2 respectively.

Model 1
\[
h = \left(\frac{46.772}{1+\exp\left(-0.114 \times (age - 17.085)\right)}\right) \times 0.908 \times \sqrt{\frac{topheight}{age}}
\]

Model 2
\[
h = \left(\frac{41.832}{1+\exp\left(-0.128 \times (age - 28.664)\right)}\right) + 11.098 \times SI
\]

For the first model (equation 12), the second explanatory variable was top height/age which entered into the equation as a multiplicative variable (equation 12). The second explanatory variable of the second model (equation 13) was SI which was a partially qualitative value and it entered into the model in additive manner. The distributions of the predictions done by those two models separately for each site class are given in the Figure 4.
3.6 Finally selected models to predict height

When observing Figure 4a, the differences in the shapes of the curves were clearly highlighted. For the model 1 (equation 12), the constructed curves tend to come together at lower ages which could biologically be well explained since, for trees growing on different sites, growth starts differentiating at a young age, not in very early stages. Moreover, the growth predicted by the models slowed down after 30 years of age, which is very much compatible with the observed rose gum growth in Sri Lanka.

Figure 4b also shows evidence that the second model departed from biological reality. Curves constructed for model 2 (equation 13) for height were parallel to each other in each case. At very low ages, i.e., 5 years, the growth differences were very high and equal for all three classes. This phenomenon would not occur in reality and therefore could not be explained biologically.
The reason for the different behaviour of the two different models (equations 12 and 13) was due to the way of entering the second explanatory variable to the model. In equation 12, this variable entered in multiplicative manner and therefore the growth rates were not parallel at different stages of the growth (Figure 4a). However, the second variable entered into equation 13 in additive manner and therefore it maintained a parallel relationship for the growth of dbh and height in different site classes as shown in Figure 4b.

Therefore, the second model was impossible to explain biologically and moreover, its predictions were unachievable at the early stages. Therefore the model given in equation 12 was finally selected to predict the height for all the classes of rose gum plantations in Sri Lanka. The results of the validation of the finally selected model predictions with the raw data are given in Figure 5.
According to the results of the validation of the finally selected model, as given in Figure 6, height values at age 30 (plantation ID: BK 2,10) indicated a slightly lower value even for the site Class III. This could be due to very poor site quality in the region for rose gum growth. Other than this particular age, the rest of the data were well covered with the selected models suggesting suitability for field use.

4 Discussion

One of the objectives of the present study was to construct the best models to predict height of rose gum growing as monocultures in Sri Lanka. For this purpose, all the possible combinations of selected candidate variables were tested with different transformations (altogether 60 different model structures were tested at stage three). For these combinations, exponential or logistic functions were modified using age and a site index as explanatory variables. Lee et al., (2004) did a similar procedure and a range of models with exponential and power functions was evaluated, using a variety of combinations of independent variables. Moreover, in their modelling process, Calama et al., (2003) and Adame et al., (2008) emphasised the importance of data chosen for fitting the different functions containing all the possible combinations of variables. Fernandez and Norero (2006) performed 45 linear regressions in modelling the growth of branches of individual trees in each site, management and type of branch. Therefore it is not an uncommon exercise to test the different combinations of selected variables to in forest growth modelling.

In addition to the untransformed variables, four simple transformations (logarithmic, square root, square and inverse) of the variables were used in this study on the assumption that those transformations can biologically be explained. However, complex transformations were sometimes used in forest growth modelling. As an example, Boisvenue et al., (2004) used arcsine and cosine transformations of selected explanatory variables to model the height growth of small trees in mixed species stands in southern British Columbia in Canada.

Age has become one of the essential explanatory variables for tree growth prediction in forestry (e.g.: Palahi et al., 2004; Adame et al., 2006; Salas and Garcia, 2006). However, according to Lee et al., (2004), although tree age is an important influencing variable on radial growth, it might simply be not available in practice. However, the attempts made by them to exclude the age from the explanatory variables to predict dbh growth of pine and oak were not successful and the resultant models indicated poor statistical performances. Therefore they decided to include age in order to obtain better results. However, for the present study, age was included from the beginning of the model construction as an essential explanatory variable. This was done with the intention of predicting height and diameter along a series of an age.

Site index models have been constructed in the past (e.g., Clutter et al., 1992; Fontes et al., 2003; Palahi et al., 2004; Louw and Scholes; 2006) mainly to classify the forest site types. Moreover, site indices have been used as explanatory variables in some growth prediction models. Pienaar and Harrison (1989) and Soares et al., (1995) used dominant height as a site representation variable to predict basal area and height respectively. However, a top height related index was selected for this study to represent the site quality since top height is known as independent from the competition (Clutter et al., 1992; Philip, 1994).

Since tree crown is responsible for the photosynthesis process, it was possible to use crown variables to indicate the competition among trees. However, these were not included as explanatory
variables in this study, because crown assessment in the field is expensive and time consuming. For the same reasons, Lee et al., (2004) eliminated the crown parameters from model construction.

A validation procedure was not followed at the model construction stage one and two of this study. The reason was the poor statistical qualities of the constructed models at those two stages. In addition to the qualitative tests, a combination of three quantitative tests, i.e., average model bias, mean absolute difference and modelling efficiency, were used for the model evaluation in this study. Rodriguez et al., (2003) suggested that the mean squared error is another good indicator of the model quality. However, following the work of Soares et al., (1995), Hein et al., (2007) and Rammig et al., (2007), the combination of the above three quantitative tests was used.

In any modelling project, several aspects of the posed problem need to be reconciled with conceptual, mathematical, engineering and ecological aspects (Huang et al., 2003). Growth equations have a quite different character from standard equations; they are heuristic generalisations rather than applications of an underlying theory (Huang et al., 2003). No single model will be best for all purposes and it is prudent not to expect otherwise. There are several reasons for this. First models are simplifications of reality and reflect the inherent inclinations, limitations, assumptions, biases and purposes of the modeller. Secondly, models contain collective knowledge gained from previous experience along with current information. As such, models are dynamic and change as new knowledge and more relevant information become available (Amateis, 2003). It is also known that increased complexity of a model reduces the generality of the considered model (Radonja et al., 2003).

References


