Comparison of selected Nordic stand growth models for Norway spruce, Scots pine and birch

Kjell Andreassen* and Bernt-Håvard Øyen


Abstract. Thirteen Nordic stand growth models have been validated by use of a test data set from long-term research plots in Norway. The evaluated data was from time-series of even-aged, pure stands of Norway spruce, Scots pine and birch (silver birch and downy birch). In selected models from Finland, Norway and Sweden measures of site productivity, mean tree size and various stand characteristics are represented. Different models display both strengths and weaknesses in their predicting ability. Several measures of precision and bias have been calculated and the models are ranked due to their performance. We observed site quality, stand density and average tree size as the three main components in the models. Basal area increment model for spruce from Sweden had the lowest standard deviation with 23%. The mean $R^2$ between residuals and stand characteristics from this model was also low (1.3%), which indicates that independent variables are well included. For Scots pine and birch, Finnish volume increment models showed the best fit to the Norwegian test data, with a $R^2$ between residuals and stand characteristics of 2.8 and 6.7%, respectively. Several of the models from Sweden and Finland predicted the growth as well as stand models frequently in use in Norway. The results indicated that similar forest conditions and traditional even-aged forest management practice in the Nordic countries could be seen as a suitable basis for developing a joint family of growth models. By careful recalibration of existing models, a reasonable accuracy could be achieved and the prediction bias could be reduced.

Key words: Growth models, accuracy, bias, Norway spruce, Scots pine, birch, evaluation, Nordic countries.

Authors’ address: Norwegian Forest and Landscape Institute, Høyskoleveien 8, N-1432 Ås, Norway; *e-mail: kjell.andreassen@skogoglandskap.no

Introduction

In Europe, traditional forest management has since the early 19th century relied on yield projection tables to estimate yields for management decisions and forest planning. Most yield tables or growth models have been constructed on basis of development in even-aged stands. Worldwide there is an extensive literature on growth modelling of such stands, less in uneven-aged stands and stands with mixtures of tree species (e.g. Vanclay, 1994; Porte & Bartelink, 2002). The first yield tables from the Nordic countries were based on graphic smoothing (e.g. Blomqvist, 1872; Stalsberg, 1882; Jonson, 1914; Møller, 1933; Brantseg, 1951). The first to apply regression analysis for smoothing and predicting of diameter growth in Nordic conifer stands was Petterson (1937) in Sweden and Eide & Langsæter (1941) in Norway. Models for estimating volume increment directly or through diameter or basal area increment have for more than 50 years been the “engine” in most yield models in the Nordic countries (for review see Vuokila & Valiaho, 1980; Ekø, 1985; Gustavsen, 1998;
Øyen, 2001; Hasenauer, 2006). There seems to be different opinions among authors about what type of models which are most suitable when predicting growth and yield. For a wide range of management alternatives and utilization options it is important to have flexible projecting systems (Hasenauer, 2006). Flexible and accurate systems, which are based on low cost inventory input data, seem desirable by most planners and foresters.

Large-scale forestry scenario models in Norway and the other Nordic countries have traditionally been based on even-aged management of the forest. Thus, area projected growth models (e.g. volume or basal area pr hectare, or the “mean tree size” combined with number of trees pr hectare) have been developed and applied in Norway (Braastad, 1974; Blingsmo, 1984; Opdahl, 1992; Øyen, 2001; Andreassen & Øyen, 2002; Nilsen & Øyen, 2002; Bollandsås et al., 2008; Andreassen & Tomter, 2008). Evaluations of some of these models have indicated rather small deviations when applied outside the range of the base material (Hobbelstad, 1980; Eid & Erikson, 1991; Pettersen, 1997; Øyen & Nes, 1997). Nevertheless, some Norwegian growth models have been criticized for giving biased prognosis (e.g. Eid, 2003; Eid & Hobbelstad, 2005; Bollandsås et al., 2008), especially when they are extrapolated in prolonged rotations.

The Nordic countries with the exception of Iceland are situated between the 55th° and 71th° latitude and the 5th° to the 30th° longitude. The forest in these countries has a dominance of Norway spruce, Scots pine, silver birch and downy birch. Growth conditions in Fennoscandia are characterized by relatively short growth seasons and cold winters. However, number of growth days (air temperature > 5°C) varies from less than 130 in the far north and in the high mountains to more than 200 in southernmost parts of Sweden and Norway. Humidity varies from very humid in the far west to very dry areas in the eastern parts or in the rain shadows. Part of the previous growth variation for the main tree species observed in the Nordic countries may be due to latitudinal or altitudinal cline, or climatic/genotype interactions. The rest seems to be attributed to differences between provenances, soil factors, stocking or silvicultural treatment (i.e. Christie & Lines, 1979). So far, most work on empirical growth models in the Nordic countries have been performed on a regional or a national scale. We have poor information whether growth models developed on the base of a defined material in one of the Nordic countries could be applied with a satisfying accuracy in other countries and regions. Models showing some bias but otherwise behave proper in different stand conditions can probably be adjusted to achieve appropriate accuracy in Norway outside the geographical range of the original model. A careful recalibration with new coefficients for the same variables will probably work better under Norwegian conditions. On such a background the objective of this paper is fourfold:

1) Describe selected growth equations and variables that are commonly applied in modelling forest growth in pure even aged stands of spruce, pine and birch in Finland, Sweden, and Norway.
2) Report the statistical and behavioural properties, accuracy and bias in some selected Nordic growth models using a comprehensive test data set from permanent research plots in Norway.
3) On the base of the observed differences present tentative conclusions about the accuracy in predictions of growth outside the geographical range of the base material.
4) Adjust or recalibrate one model of each tree species to achieve a better fit.

We hypothesize that similar forest conditions and traditional even-aged forest management practice implies that several growth models for Norway spruce, Scots pine and birch in the Nordic countries could be used directly or recalibrated.
to behave consistently and in such provide reasonable projections in flexible information systems.

Material and Methods

The test data set

The first Norwegian Forest and Landscape Institute growth and yield plot was established in 1915, and increment data from even-aged stands up to 2007 was applied in the work. The test data set comprises 3966 growth periods in Norway spruce, 3144 in Scots pine and 556 in birch stands. Mean increment period were 5 years for all three species. Data from different regions of Norway except the western parts are represented, however the majorities of the plots are located in Southeast Norway. The elevations are from sea level up to 1020 m in spruce, to 885 m in pine and to 790 m in birch. A wide variety of thinning programs are applied, however most stands have been low thinned or crown thinned (B-grade or C-grade). Unthinned (A-grade) and a few heavy thinned stands (D-grade) are also represented. The mean thinning percentage (in volume) is 16% in spruce, 17% in pine and 19% in birch. Old stands with age above 140 years in Norway spruce are poorly represented, while 3% of the plots in Scots pine are older than 140 years. The oldest stand in birch is 95 years (Table 1). To harmonise the models and the test material, all stands with age below 20 years were excluded. Mean plot size varied from 900 m$^2$ in Norway spruce to 1100 m$^2$ in Scots pine and 1400 m$^2$ in Birch. Silver birch and downy birch are gathered in one group since spontaneous mixtures of the two species occurs side by side in Norway.

Site quality in Norway is described according to the $H_{40}$ site index (SI) system (Tveite & Braastad, 1981). For converting $H_{40}$ into the Swedish $H_{100}$ system we used the following equations (Tveite, 1980):

Norway spruce:  

$$H_{100} = 1.260 \times H_{40} + 5.45 \quad (1)$$

Scots pine and birch:  

$$H_{100} = 1.351 \times H_{40} + 2.61 \quad (2)$$

The growth equations

On the background on requests to yield researchers in the Nordic countries, commonly applied national growth and yield models were identified. A large number of models are presently being utilized for different purposes, e.g. from Finland more than 40 stand increment models have been developed for various regions and tree species (Gustavsen, 1998). We selected a few models from each country including independent variables available from common inventory data (Table 2). The Norwegian models are based on other data (Strand, 1983; Solberg & Tveite, 2000) than the test data.

Table 1. Mean values for test data. SD in brackets.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Norway spruce</th>
<th>Scots pine</th>
<th>Birch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=3966)</td>
<td>(n=3144)</td>
<td>(n=556)</td>
</tr>
<tr>
<td>Elevation, m.a.s.l.</td>
<td>215 (163)</td>
<td>232 (156)</td>
<td>201 (171)</td>
</tr>
<tr>
<td>Plot size, m$^2$</td>
<td>870 (557)</td>
<td>1090 (833)</td>
<td>1440 (748)</td>
</tr>
<tr>
<td>Mean stand age at breast height ($T_{1.3}$), yrs</td>
<td>40 (20)</td>
<td>52 (32)</td>
<td>40 (14)</td>
</tr>
<tr>
<td>Site index, mean top height at age 40 yrs ($H_{40}$), m</td>
<td>18.4 (4)</td>
<td>13.7 (3)</td>
<td>15.5 (5)</td>
</tr>
<tr>
<td>Basal area (G3), m$^2$ ha$^{-1}$ of trees with d $\geq$ 25 mm</td>
<td>29.4 (10)</td>
<td>21.4 (8)</td>
<td>17.2 (7)</td>
</tr>
<tr>
<td>No stems (N3) pr ha$^{-1}$ with dbh $\geq$ 25 mm</td>
<td>1636 (1062)</td>
<td>1555 (1322)</td>
<td>1967 (1924)</td>
</tr>
<tr>
<td>Standing volume (V3) m$^3$ ha$^{-1}$</td>
<td>232.3 (127)</td>
<td>145.4 (76)</td>
<td>117.2 (73)</td>
</tr>
<tr>
<td>Basal area mean diameter (D3), cm</td>
<td>16.8 (5.4)</td>
<td>15.1 (5.2)</td>
<td>13.0 (5.3)</td>
</tr>
<tr>
<td>Lorey’s mean height ($H_L$), m</td>
<td>15.0 (4.7)</td>
<td>12.9 (3.9)</td>
<td>13.9 (4.9)</td>
</tr>
<tr>
<td>Volume increment (iV), m$^3$ ha$^{-1}$ yr$^{-1}$</td>
<td>13.5 (13)</td>
<td>6.4 (3)</td>
<td>5.4 (3)</td>
</tr>
<tr>
<td>Basal area increment (iG), m$^2$ ha$^{-1}$ yr$^{-1}$</td>
<td>1.07 (0.43)</td>
<td>0.66 (0.31)</td>
<td>0.60 (0.28)</td>
</tr>
<tr>
<td>Diameter increment (iD), mm</td>
<td>2.9 (1.3)</td>
<td>2.2 (1.1)</td>
<td>2.2 (1.1)</td>
</tr>
<tr>
<td>Volume increment percentage (pV), %</td>
<td>6.7 (3.7)</td>
<td>5.6 (3.4)</td>
<td>5.4 (3.1)</td>
</tr>
</tbody>
</table>
Table 2. Test models used. FG, GG, LG denotes % basal area of Scots pine, Norway spruce and birch, respective. GK is % other tree species. G1, G2 and G3 is basal area before thinning, removal and after thinning, respective. Ho is top height. TT is total stand age. Other variables - see Table 1.

<table>
<thead>
<tr>
<th>No.</th>
<th>Growth model</th>
<th>Ref.</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \ln iG = -1.3933 - 0.2309/1000*G3 + 0.67033* \ln G3 + 0.23347* \ln N3 - 0.34588* \ln T )</td>
<td>Aggestam, 1985</td>
<td>S</td>
</tr>
<tr>
<td>2</td>
<td>( \ln iV = -1.28925 -0.00831<em>Ho +0.91681</em> \ln Ho+0.03879* \ln G%-0.00479* \ln N1-0.03662<em>N1/1000+0.01373</em>G1+0.15359* \ln G1 )</td>
<td>Eriksson, 1985</td>
<td>S</td>
</tr>
<tr>
<td>3</td>
<td>( \ln pV_5=9.25266-1.99553* \ln TT )</td>
<td>Gustavsen, 1998</td>
<td>F</td>
</tr>
<tr>
<td>4</td>
<td>( \ln iD = 4.68 + 0.205* \ln H_{100} -0.206* \ln N3-0.783* \ln T_{1,3} )</td>
<td>Solberg &amp; Tveite, 2000</td>
<td>N</td>
</tr>
<tr>
<td>5</td>
<td>( \ln iG = -2.0065 – 0.1328/1000<em>G3 + 0.42115</em> \ln G3 +0.37398* \ln N3 + c<em>T_{1.3} -0.19424</em> \ln T_{1.3} -0.2130/1000*GK )</td>
<td>Aggestam, 1985</td>
<td>S</td>
</tr>
<tr>
<td>6</td>
<td>( \ln iV = -3.24142 -0.00115<em>Ho -0.40621</em> \ln Ho+0.05757<em>H_{100}-0.21673</em> \ln G_{100} )</td>
<td>Eriksson, 1985</td>
<td>S</td>
</tr>
<tr>
<td>7</td>
<td>( \ln pV_5=2.9509+1.0599*(5.69-\ln TT) )</td>
<td>Gustavsen, 1998</td>
<td>F</td>
</tr>
<tr>
<td>8</td>
<td>( \ln iG =-1.13493+0.40654* \ln G3 – 0.63555* \ln T_{1,3} +0.54003* \ln H_{100} )</td>
<td>Persson, 1992</td>
<td>S</td>
</tr>
<tr>
<td>9</td>
<td>( \ln iG = 0.4298 + 0.63464* \ln G3 + 0.17829* \ln N3 -0.73932* \ln T_{1,3} +0.2557/1000<em>H_{100} -0.2670/1000</em>GK )</td>
<td>Aggestam, 1985</td>
<td>S</td>
</tr>
<tr>
<td>10</td>
<td>( \ln iG = -0.73932* \ln T_{1,3} +0.2557/1000<em>H_{100} -0.2670/1000</em>GK )</td>
<td>Aggestam, 1985</td>
<td>S</td>
</tr>
<tr>
<td>11</td>
<td>( \ln iG = -0.73932* \ln T_{1,3} +0.2557/1000<em>H_{100} -0.2670/1000</em>GK )</td>
<td>Aggestam, 1985</td>
<td>S</td>
</tr>
</tbody>
</table>

The thirteen equations has four different response variables; mean tree diameter increment (iD) in mm pr year, basal area increment (iG) in m² pr year, stand volume increment in m³ pr year (iV) and volume increment in percent of standing volume over five years (pV₅).

**Statistical methods and test procedures**

Thirteen growth models were compared by their ability to explain variation and by visual inspection and statistical tests of the residuals. Mean differences between predicted and observed increment were evaluated. The corresponding standard deviations were also addressed. Let \( y_i \) denote the observed and \( \hat{y}_i \) the predicted increment for observation \( i \), \( i = 1, 2, ..., n \). The model performance was computed as the average model bias (MB), and the corresponding model bias percent (MBp) was computed by:

\[
MB = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{n} \quad (3)
\]

\[
MBp = \frac{100 \times \sum_{i=1}^{n} (y_i - \hat{y}_i)}{n} \quad (4)
\]

The mean absolute difference (MA) was calculated by the formula of Burk (1986):

\[
MA = \left( \sum_{i=1}^{n} | y_i - \hat{y}_i | \right) / n \quad (5)
\]

Finally, model efficiency was computed by:

\[
EF = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (6)
\]
Model efficiency (EF) also compares predicted growth with observed value where zero or negative EF values indicate large deviations and a poor model (Vanclay & Skovsgaard, 1997). A good fit is indicated by an EF-value close to 1.

We also used the least squared method to estimate the squared correlation ($R^2$) and the F-value for identifying patterns in the comparison of observed and predicted growth (SAS 2004). In addition, the standard deviation (SD) and skewness were estimated. To illustrate and compare the patterns and performance of the models in figures we used a spline procedure (SAS 2011) to smooth a curve of the deviations between estimated and observed growth.

Results

In the following, general results are presented, and then results of each tree species and each growth model. The results of the evaluation procedures are presented in Table 3–4 and Figure 1. The models predicted large variations in growth estimates. Largest deviations were found in stands with extreme values of the input variables, and for birch models (cf. Fig. 1). With four different response variables in the models, we emphasized relative deviations rather than absolute values as most suitable for comparisons.

The iD-models provide the smallest deviation in average increment with 1–13% relative model bias (MBp). The iV-models provide less than 11% deviation, the iG-models 17–36%, and the pV-models 6–66% deviation. For birch models the standard deviation of MBp is considerable higher than for spruce and pine. The mean absolute difference (MA) for the iD-models varies from 0.6–0.7 mm yr$^{-1}$, the iG-models from 0.16–0.32 m$^2$ yr$^{-1}$, the iV-models are in the range from 1.75 to 2.41 m$^3$ ha$^{-1}$ yr$^{-1}$ and the pV-models show differences from 1.1 to 2.7% volume increment. However, MA is an absolute value and is comparable only for models with the same dependent variable.

The skewness is based on the deviation predicted minus observed growth and show values from −1.3 to +0.9 (Table 3). Since the skewness formula cubes the deviations, large deviations are emphasized and the sign of the mean deviation is maintained. Skewness indicates the tendency of the deviations to be larger in one direction than another. All the iD-models predict negative skewness which indicates an underestimation. Other types of models have both negative and positive skewness. However, means of absolute devia-

<table>
<thead>
<tr>
<th>Model no.</th>
<th>Response variable</th>
<th>MBp (SD)</th>
<th>Rank of MBp</th>
<th>SK</th>
<th>MA</th>
<th>EF</th>
<th>Rank of EF</th>
<th>F-value</th>
<th>Rank of F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>iG</td>
<td>79 (23)</td>
<td>5</td>
<td>-0.80</td>
<td>0.32</td>
<td>-3.11</td>
<td>5</td>
<td>3892</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>N. iV</td>
<td>100 (26)</td>
<td>1</td>
<td>0.10</td>
<td>2.41</td>
<td>0.59</td>
<td>2</td>
<td>7800</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>spruce iV</td>
<td>94 (25)</td>
<td>4</td>
<td>-0.70</td>
<td>1.38</td>
<td>0.63</td>
<td>1</td>
<td>12180</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>iD</td>
<td>101 (28)</td>
<td>2</td>
<td>-0.68</td>
<td>0.60</td>
<td>0.03</td>
<td>3</td>
<td>5029</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>iD</td>
<td>98 (33)</td>
<td>3</td>
<td>-0.51</td>
<td>0.69</td>
<td>-0.06</td>
<td>4</td>
<td>3021</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>iG</td>
<td>76 (27)</td>
<td>5</td>
<td>-0.98</td>
<td>0.24</td>
<td>-1.80</td>
<td>5</td>
<td>2120</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>S. iV</td>
<td>89 (34)</td>
<td>2</td>
<td>-0.78</td>
<td>1.75</td>
<td>-0.20</td>
<td>2</td>
<td>3145</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>pine iG</td>
<td>83 (24)</td>
<td>4</td>
<td>-1.32</td>
<td>0.19</td>
<td>-1.01</td>
<td>4</td>
<td>3597</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>pV</td>
<td>110 (32)</td>
<td>1</td>
<td>-0.61</td>
<td>1.14</td>
<td>0.68</td>
<td>1</td>
<td>8256</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>iD</td>
<td>87 (33)</td>
<td>3</td>
<td>-1.10</td>
<td>0.68</td>
<td>-0.45</td>
<td>3</td>
<td>2095</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>iG</td>
<td>138 (46)</td>
<td>2</td>
<td>0.94</td>
<td>0.23</td>
<td>0.33</td>
<td>1</td>
<td>365</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>Birch pV</td>
<td>166 (62)</td>
<td>3</td>
<td>0.72</td>
<td>2.66</td>
<td>0.13</td>
<td>2</td>
<td>812</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>iD</td>
<td>105 (42)</td>
<td>1</td>
<td>-0.60</td>
<td>0.62</td>
<td>-0.70</td>
<td>3</td>
<td>400</td>
<td>2</td>
</tr>
</tbody>
</table>

K. Andreassen and B.-H. Øyen
tions often differ from means of percentage values since larger values are emphasized in absolute deviations.

Simple F-tests provide a good basis for comparison where the same tree-species are used. Using the least squared method for calculating the F-value, spruce models have the highest values and best fit (Table 3). The lowest F-value is found for the birch models. In a linear regression between observed and predicted values, an intercept close to 0 and a slope close to 1 indicate a good fit (Mayer et al., 1994).

Model efficiency (EF) range from –3.1 to +0.7 (Table 3). Seven of the models have negative values which indicate a poor fit.

We examined how the models behave under different stand and growth factors by plotting the residuals against the four selected variables site index, standing volume, age and mean tree diameter (Fig. 1). To be able to compare different models in the same figure, we decided to use a smoothed curve of the deviations.

The Person correlation matrix between the residuals and main variables indicate the influence and how good variables are included in the model (Table 4). A low correlation indicates that those variables are already well included in the model and cause only small deviations. Mean correlations are used for ranking the models, and the most important factor is tree size, expressed by $H_1$ and $Dg$ explaining 11 and 10% of the deviation, respectively. The second most important factor is site index ($H_{40}$), explaining 9%. Other factors like standing volume ($V_3$) or age ($T_{1.3}$) account for less than 8% of the deviation. The two density factors basal area ($G_3$) and spacing index ($S%$) explained 5.5% and 4%. In this test the Norwegian models of Strand (1983) showed the lowest fit. Especially the size variables are responsible for a large proportion of the deviation in these models.

In addition, new coefficients for three models are proposed to achieve more accurate growth prognosis under Norwegian conditions (Table 5; Fig. 2).

Norway spruce

For Norway spruce, models 2, 5 and 6 have a small average model bias (MBp) (Table 3). Model 1 and 3 generally underestimate the growth. The SD (standard deviation) is smallest for model 1 and largest for model 6, and range from 23 to 33%. The mean absolute difference (MA) has smallest value for model 1 and highest value for model 2. Skewness is 0.1 as the best score (model 2).
Figure 1. Residuals (in % of observed) in different sites and stand properties. Four upper boxes are Norway spruce models 1–6, four medium boxes are Scots pine models 7–11, and four lower boxes are birch models 12–14.
while the other spruce models have about the same skewness. The highest F-value is identified for model 3 in this test. The general high F-values for Norway spruce indicate a combination of good fit and a high number of observations. Model 2 and 3 show best performances in the model efficiency test (EF), while model 1, 5 and 6 indicate a poor fit.

Figure 1 illustrates well how the model behaves in different stand conditions. It is interesting that the predicted growth is almost constantly underestimated with about 20% for model 1. We experienced that these four main variables (SI, V3, Age and Dg) were well included in model 1 since the residuals show almost straight lines, in spite of the more or less constant “intercept” bias. This is also underlined in table 4 where model 1 show best fit. Model 2 is clearly biased in the outer part of the test data, but still ends up with an average of predicted growth equal to the observed growth.

The correlation matrix indicates large error for site index in model 3, while stand age is well included for this model (Table 4). The mean of these individual \( R^2 \) values for all this important stand characteristics indicates well the performance with the overall lowest mean for model 1. Model 2 and 5 display also low \( R^2 \), however 3–4 times worse than model 1. Model 6 is identified with the highest \( R^2 \) value and indicates low performance.

**Scots pine**
The relative model bias (MBp) for Scots pine is highest for model 7 with an average underestimation of 24%, and smallest bias for model 10 with an overestimation of about 10% (Table 3). The SD of MBp range from 24 to 34% with smallest value for model 9 and largest value for model 8. Skewness is negative for all models and range from \(-1.3\) to \(-0.6\) and indicates tail of the underestimation. Largest skewness is identified for model 9, while model 10 has the smallest skewness. Model 10 has the highest score of F-value with 8256, while model 7 and 11 address almost four times lower scores (Table 3). The correlations between residuals and stand characteristics are in general highest for model 9 and 12 which indicate low performance (Table 4). Smallest correlation between residuals and stand characteristics is found for model 10, except for site index (Table 4). The visualisation of the residuals of the models (Fig. 1), underlines the results that model 10 show best performances, except site indices which is not well include in the model. The lack of the variable site index in model 10 is probably the reason for low performance if we analyse the residuals versus site index. The figure also illustrates low performances for model 8 and 11, which is in line with the results in table 3 and 4.

**Birch**
Model 14 has the smallest model bias (MBp), SD and skewness, while model 13 has the highest values of MBp and SD (Table 3). The mean predicted growth for model 14 is only 5% higher than the observed mean, while model 12 and 13 overestimate the growth with 38 and 66%, respectively. However, considering the correlation between observed and predicted growth, almost opposite result appears with poor F-values for model 12 and 14, and highest F-value for model 13 (Table 3). The mean absolute difference (MA) is smallest for model 12 and poorest for model 13. Model efficiency (EF) indicate low performance for model 14, and medium performance for model 12 and 13 (Table 3). The correlations between residuals and stand characteristics indicate well performance for model 13 except site index which is not so well included in the model (Table 4). The \( R^2 \) values for model 12 and 14 indicate low performance for some variables, but good performance for other variables. Model 13 has the best \( R^2 \) ranking of the birch models. Figure 1 illustrates also well how the models behave in different stand conditions, e.g. model 13 shows smallest deviation of the birch models. All models
have certain shortcomings especially pronounced for model 14. For model 12, especially the volume and basal area variables seem not to be well included.

**New coefficients**

We used the same variables as in the original models to calculate new parameter estimates for one of the best models for each tree species (model 1, 10 and 13) based on our test data (Table 5). These coefficients changed a bit compared with the original models. The average model bias (MBp) is therefore almost harmonized to zero in this adjustment to the test data. Consequently the performance of these models in different growth and stand conditions is indicated with a more harmonised curve around the zero line with smaller residuals (Fig. 2). However, the new models have the same weakness as the original models and the residual curves behave almost the same way as these (cf. Figs. 1–2).

**Discussion**

In this study, we experienced that the selected Nordic stand growth models work well at medium site indices, and with medium stand properties concerning stand volume, mean tree size and stand age. The evaluation revealed several weaknesses for the selected growth models, particularly in the margins of our test data. This also corresponds with similar studies (e.g. Hasenauer, 2006; Bollandsås & Næsset, 2009; Härkönen, 2010). Also the Norwegian models we tested, and not based on the test data, displayed the same or even worse behaviour compared to Finnish and Swedish models.

Several interesting patterns were identified. The iD-models from Norway do not show any better fit than models from the other Nordic countries. The model 5 is based on 570 Norwegian monitoring plots in pure spruce stands measured from 1991 to 1996 (Solberg & Tveite, 2000), while models 6, 11 and 14 are based on 200 NFI plots in mixed stands measured from 1971 to 1978 (Strand, 1983). Even though such models give reasonable predictions in medium site indices, and at medium tree size and stands, they do not work very well outside average conditions. The result indicates that models based on rather few stands and short-term investigations must be looked upon as inappropriate management tools.

Another set of Norwegian stand models based on separate NFI data also revealed some weaknesses in evaluation with our test data (Andreassen & Tomter, 2008).
Residuals can also be ascribed to different climate, soil, vegetation type and genetic components (Stage & Wykoff, 1993). For spruce the Swedish model 1 has the largest mean deviation and the poorest model efficiency, but shows much better fit in other test procedures. Similar with model 6, this model was originally developed for prognosis in mixed stands (Agestam, 1985). For Scots pine model 10 has the second lowest squared correlation coefficient and the highest model efficiency. None of the three birch models give accurate predictions, but model 13 (Gustavsen, 1998) displays the best performance over a wide range of stand and site conditions.

The test data represent all regions of Norway except the western region, and includes several thinning regimes and a wide range of sites and stand characteristics are present. Our test data is measured over a long period, and there were different goals for these trials. Most of the plots established before 1950 originate from natural regeneration and plots after 1950 are mostly planted, in line with NFI data (Larsson & Hylen, 2007). The test data cover well Norwegian forest areas. However, normally NFI data indicate less growth than well treated experimental data at equal site quality (Andreassen et al., 2008) and this effect might identify some of the observed bias.

Most of the growth equations have been constructed on the basis of the ordinary least square method (linear regression) and transformation and combinations of the explanatory variables specified. In ten of thirteen models logarithmic transformations have been made, only Strand (1983) has based his work on additive models. In the construction phase several challenges are subjected to how to deal with heteroscedastic variance, non-normal distributed errors and independent explanatory variables (Clutter et al., 1983). Independency between variables is important to avoid collinearity diagnostics (Chatterjee & Price, 1977). How these elements are handled vary from author to author. In Swedish growth models, the $H_{100}$ site index system is used and this is converted (Tveite, 1980)

![Figure 2. Residuals (in % of observed) of growth models with new coefficients in different sites and stand properties. M16 = Agestam (1985) for Norway spruce, M17 = Gustavsen (1998) for Scots pine, M18 = Gustavsen (1998) for birch.](image-url)
to the $H_{40}$ (Tveite & Braastad, 1981). Such linear conversion might involve bias, especially in the periphery of the data coverage (i.e. Blingsmo, 1988; Eid & Eriksson, 1991; Pettersen, 1997). Tveite (1980) indicates only small deviations when a linear conversion is used. However, in the Finnish models of Gustavsen (1998) the site quality is included with the explanatory variables top height and total age.

It is not straight forward to make meaningful comparisons between alternative models. All the chosen Nordic models are practical orientated growth models basically constructed for schedule timing and yield of a timber harvest, and they rely on common inventory data. To find an appropriate menu of diagnostic tests is an important issue, and among many we have made a limited selection. Vanclay & Skovsgaard (1997) propose some methods for evaluation forest growth models and several of these are included in our investigation. Oderwald & Hans (1993) emphasize the importance of logical and biological components in models like sign of coefficients, expected variables, reasonable transformations and expected values in extremes. Reynolds (1984) and Burk (1986) suggest different methods to estimate the errors in model predictions as model bias (MB) and mean absolute difference (MA) which is in line with our methods. Sensitivity analysis is a frequent used evaluation method (e.g. Jørgensen, 1986), but is omitted in our investigation.

Different response variables in the models make comparison difficult, even though we used the relative bias (MBp) in the evaluation. E.g. since volume growth normally is an important goal of forest prognosis, growth via basal area- or diameter increment models also includes uncertainties with tree height estimation which probably increase the bias of volume growth even more (Clutter et al., 1983).

The explanatory variables should be carefully chosen to ensure biologically realistic predictions across the whole range of possible conditions. The independent variables in the models seem reasonable. They are found to be of the three major types: site factors (e.g. site index, top height, age), mean tree size (e.g. mean tree height, mean tree diameter) and competition/stand descriptions (e.g. basal tree height, standing volume, stems ha$^{-1}$). This is in line with suggestions of other authors (Hasenauer, 2006; Clutter et al., 1983), and seems to include the most important biological and logical components (Oderwald & Hans, 1990). Growth models including several variables are generally more adapted to the basic data than models with few variables (Hasenauer, 2006). Such models are often sensitive in the outer part of the basic data, especially if same variable is repeated or including power terms (Vanclay, 1994). Including related parameters several times will often lead to a better fit, but is rarely a good option in models used for extrapolation beyond basic data (Rennolls, 1995). However, to study specific management practices, it is interesting to also include significant variables reflecting special issues (Andreassen & Tomter, 2003). We emphasize that our evaluation reflects common forest condition without extrapolation since only real observed data is used. We experienced the performances of models including several variables (e.g. Strand, 1983; Eriksson, 1985; Agestam, 1985; Persson, 1992) were not much better than models including few variables (e.g. Gustavsen, 1998). The reasonability of the signs of the coefficients are difficult to evaluate since some variables are included several times in the model, some of these have opposite signs, some variables are transformed, and some transformations are sophisticated expressions with several variables included. Errors in measuring and calculating forest variables can also address several sources of deviations in growth analysis (c.f. McRoberts et al., 1995; Kangas & Kangas, 1999).

The new coefficients of the three models seem reasonable and imply a careful adjustment with the original models of Agestam
(1985) and Gustavsen (1998). It is interesting to notice that the correlation is better and the variance lower than in Norwegian models developed from the same data (e.g. Braastad, 1975, 1977, 1980; Blingsmo, 1984, 1988). However, our test data set is larger than the corresponding basic data.

Conclusions

Success in developing models depends on selecting the suitable low cost high quality variables, collecting good data, using reliable coefficients estimation and by carefully evaluation procedures. Our work has shown that various Nordic models have strengths and weaknesses. It is difficult to specify which models that show best performance since different test procedures show different results. However, for Norway spruce model 1 (Agestam, 1985) works well over the most range of test data in spite of an almost linear bias. The Scots pine and birch models 10 and 13 (Gustavsen, 1998) do also show an adequate performance when the Norwegian test data set was applied. For providing useful overall predictions in Norwegian conditions we suggest that models 1, 10 and 13 are further recalibrated, and based on the test data set from Norway we have presented recalibrated models.

References


Burk, T.E. 1986. Growth and yield model validation: Have you ever met one that you liked? – Allen, A., Conney, T.C. (eds.) Data management issues in...
Møller, C.M. 1933. Bonitet, vekst og produksjon hos ospe i Sør-Norge. (Site index, growth and production in aspen (Populus tremula L.) in South-Norway). – Communications from Norwegian Forest Research Institute, 6/02, 7–9. (In Norwegian).
Comparison of selected Nordic stand growth models for Norway spruce, Scots pine and birch

Institute, 47/18, 24 pp. (In Norwegian).
Tveite, B. 1980. Samband mellom svensk boniteringssystem (H100) og norsk boniteringssystem (H40). (Relationship between Swedish site index system (H100) and Norwegian site index system H40). Internal note, Norwegian Forest Research Institute, 17/1-1980. 2 pp. (In Norwegian).

Received February 10, 2012, revised March 09, 2012, accepted March 27, 2012