CHAPTER 6

GENERALIZING SAMPLE TREE INFORMATION

JUHA LAPPI, LAURI MEHTÄTALO AND KARI T. KORHONEN
Finnish Forest Research Institute

Some characteristics, e.g. tree height, may be rather expensive to measure, so that we may not want to measure them for all trees in an inventory but just for a subset, called here *sample trees*. We will then want to generalize the sample tree information to cover the *tally trees,* for which basic measurements, usually diameter at breast height and tree species, are available. In addition, we are often interested in many characteristics of trees which are too expensive to measure even on sample trees, e.g. stem volume or tree biomass. If these variables can be predicted using sample tree variables, we may then want to generalize the sample tree information to obtain predictions for the tally trees as well.

The first section of this chapter describes briefly different approaches to modelling sample tree variables using tally tree variables (called tally tree regression), while the second shows how auxiliary data can be utilized if the inventory data is too limited for performing tally tree regressions. The difference between real sample tree variables (e.g. tree height) and predicted sample tree variables (e.g. tree volume and biomass) will be ignored in the first two subsections, but in the third subsection we will take a closer look at the consequences of the three-level model structure created by observed tally tree variables, observed sample tree variables and predicted sample tree variables.

6.1 ESTIMATION OF TALLY TREE REGRESSION

Let y denote the sample tree variables (e,g, height, predicted volume or predicted biomass) and let $\mathbf{x}$ denote tally tree measurements (e.g. dbh). The tally tree regression model is

A. Kangas and M. Maltamo (eds.), Forest Inventory – Methodology and Applications, 85–106.

$$y_i = E\left(y|\mathbf{x}_i\right) + e_i = f\left(\mathbf{x}_i|\boldsymbol{\beta}\right) + e_i ,$$

where y_i, $\mathbf{x}_i$ and e_i are the sample tree variable, the vector of tally tree variables and the residual, respectively, for sample tree i and $\boldsymbol{\beta}$ is a vector of parameters. The assumptions regarding the form of $f\left(\mathbf{x}_i|\boldsymbol{\beta}\right)$ and the properties of e_i depend on the modelling situation, see Weisberg (1985) and Davidian and Giltinan (1995) for a general formulation of a regression model.

In forestry applications, a transformation of y is often modelled instead of y itself, in order to solve the problem of heteroscedastic residuals. In many cases, it also makes the model linear with respect to some transformations of $\mathbf{x}$. Techniques for fitting non-linear models and estimating variance functions have developed rapidly in recent years, and these are now commonly available in modern statistical software packages (e.g. Pinheiro and Bates 2000). Modelling of the dependence of predicted sample tree volume on tree diameter using non-linear regression and a variance function is demonstrated in Example 6.1.

Forest inventory data are usually collected from several plots with several sample trees per plot. Hierarchical datasets of this kind can naturally be modelled through a mixed modelling approach. A mixed model for variable y in the case of tree i on plot k is defined as

$$y_{ki} = E\left(y|\mathbf{x}_{ki,}\mathbf{b}_k\right) + e_{ki} = f\left(\mathbf{x}_{ki}|\boldsymbol{\beta},\mathbf{b}_k\right) + e_{ki} ,$$

where y_{ki}, $\mathbf{x}_{ki}$ and e_{ki} are the sample tree variable, the tally tree variable vector and the residual, respectively, for sample tree i on plot k; $\boldsymbol{\beta}$ is the vector of fixed parameters and $\mathbf{b}_k$ the vector of random parameters on plot k. Thus the random parameters vary from plot to plot. The expectations for these parameters are usually assumed to be 0 and their variances are estimated at the model fitting stage. The realized values of the random parameters can be predicted for each plot in the dataset. When using the model for prediction purposes, one can use either expectations or predictions for the random parameters, the former approach giving population-level predictions and the latter plot-level predictions. In addition to the plots, stands, clusters or measurement occasions may create additional levels of grouping in the data. For a formal presentation of the approach, readers are referred to Lappi (1993), Davidian and Giltinan (1995), Pinheiro and Bates (2000) and McCulloch and Searle (2001). Forestry applications of linear mixed models have been reported by Lappi (1991), Gregoire et al. (1995) and Mehtätalo (2004), for example, and non-linear models by Gregoire and Schabenberger (1996) and Fang et al. (2001). Both a non-linear and a linear mixed-effects modelling approach to the dataset of Example 6.1 are illustrated in Example 6.2.

An important point to note, as discussed by Diggle et al. (2002), for instance, is that in an inventory we are estimating a cross-sectional regression function (population average) for y. The resulting function is not generally a logical

one for describing how y in a single tree is related to the development of $\mathbf{x}$ over time (subject-specific curve). For instance, in cross-sectional data $H(D)$ often decreases with respect to D when D is large, even though the height of each single tree grows as D grows. This results from the fact that in inventory data trees with a large D are usually open-growing ones that have been allocating resources to diameter growth rather than height growth. Thus the growth allometry of individual trees does not provide theoretical reasons for any specific functional form for $E(y|\mathbf{x})$.

There are also approaches that do not require definition of the functional form of $E(y|\mathbf{x})$. One approach is to compute class means for y in various classes of $\mathbf{x}$. The y values for tally trees can then be obtained either by using class means directly or by interpolating them. The bias in the estimated $E(y|\mathbf{x})$ is small in these methods, but the estimation variances can be large because the large number of parameters (class means) has to be estimated. This approach does not make proper use of the smoothness of $E(y|\mathbf{x})$. There are many methods nowadays for estimating flexible regression functions without any theoretical parametric models, e.g. regression splines, kernel smoothing, local regression and smoothing splines (see Hastie et al. 2001). Smoothing splines may provide a good general basis for future large-scale inventories.

The approaches presented above may be problematic if there is more than one variable that should be generalized. In such a case, regression models are required for each variable and it is difficult to ensure that the estimates for a single tally trees are logical and harmonized. One solution for avoiding such problems is non-parametric nearest neighbour (k-nn) estimation. The principle is simple. For each tally tree we search for the sample trees that are most similar to it and calculate the (weighted) mean of y for these trees. Similarity is measured with respect to tally tree variables $\mathbf{x}$. For example, the estimate of stem volume (or saw log volume) for a tally tree is the (weighted) mean value of stem volumes (or saw log volume) of its nearest neighbours. This technique was employed by Korhonen and Kangas (1997) to generalize sample tree information for tally trees.

Example 6.1 Estimation of tally tree regression using weighted non-linear least squares

This example uses non-linear regression to fit a volume model to sample tree data. Note that the volume modelled is not the actual volume but a prediction based on diameter and height (see section 6.3). The dataset consists of 385 Scots pine trees measured on 16 plots, being a subset of the INKA data originally collected for growth and yield studies (Gustavsen 1998). All these trees were measured for diameter and height, but just 61 trees were selected as sample trees for this example (Table 6.1), the remaining 324 trees being left as tally trees.

The volume of tree i in the sample tree data seemed to follow a non-linear model

$$y_i = \exp(a + b \ln x_i) + e_i ,$$

where a and b are model parameters, x_i is the diameter of tree i, $i=1,...,N$, and N is the total number of trees in the data. Since, as seen in Figure 6.1a, the variance of the volume clearly increased as a function of diameter, it was assumed that

$$\mathrm{var}(e_i) = \sigma^2 x_i^{2\delta} .$$

Using this assumption, the model was fitted using weighted non-linear least squares. The initial guesses for the estimates, obtained from a linearized ordinary least squares regression, were a = -2.373 and b = 2.632, and the final parameter estimates obtained were

a = -2.365, b = 2.631, σ^2 = 0.01025^2 and δ = 2.677.

The variance function seemed to homogenize the variance well (Figure 6.1b), and the resulting model (see Figure 6.3) was used to predict the volumes of tally trees, giving a RMSE for the predictions of 19.04 dm^3.

Table 6.1 The sample tree data of Examples 6.1-6.3. Volume is a prediction based on diameter and height (Laasasenaho 1982). Plot 51 was not included in the modelling data.

plot	Diameter,	Height,	Volume,	plot	Diameter,	Height,	Volume,
1	5	4.9	6.17	10	14.9	14	122.60
1	8.3	7.7	23.31	11	10.3	8.4	38.33
1	9.7	9	35.99	11	12.2	10.9	66.48
2	5.2	5.6	7.28	11	5.7	7.3	10.58
2	6.3	5.8	10.93	11	7.7	9.8	24.41
2	7.6	7.2	18.58	11	7.6	8.5	21.16
2	10.1	9.1	39.34	11	10.2	10.4	44.83
2	8.1	7.2	21.09	11	9.2	10.3	36.24
3	15.2	11.6	108.14	12	12.7	13	83.84
3	10.7	9.3	44.90	12	11.6	13	70.10
3	8	7.2	20.58	12	25.7	18.6	458.78
4	12	13.1	75.47	12	12.2	13.3	79.02
4	12.7	12.1	78.74	12	13.4	13.6	96.97
4	15.1	13.4	121.07	12	19.5	17.5	253.89
5	18.7	19.6	259.58	12	22.1	19.7	361.11
5	22.2	18.9	350.64	12	18.4	18	232.59
6	3.3	3.1	2.19	13	9.1	7.8	28.27

6	5	3.9	5.45	13	10.3	8.1	37.23
6	3.7	3.8	2.95	13	5.7	5.3	8.43
6	3.1	3.1	1.93	13	9.1	7.7	27.99
6	4	3.3	3.27	13	12.1	8.7	54.22
6	4.5	3.6	4.27	14	6.7	6.2	12.94
7	13.1	10.3	72.92	14	7	5.6	13.16
7	13.5	10	75.47	14	4.9	5.6	6.46
7	15.1	10.3	96.45	15	5.4	8.2	10.41
7	9.5	7.8	30.79	15	6.5	8.2	15.06
8	21.7	14.1	257.33	16	4.8	6.5	6.89
8	10.4	9.2	42.06	16	7.4	7	17.25
8	17	13.8	156.79	16	5	5.3	6.49
8	6	7.4	11.85				
8	14.9	11.4	102.43	51	3.9	4	3.36
9	21.5	16.9	297.46	51	8.1	5.6	17.60

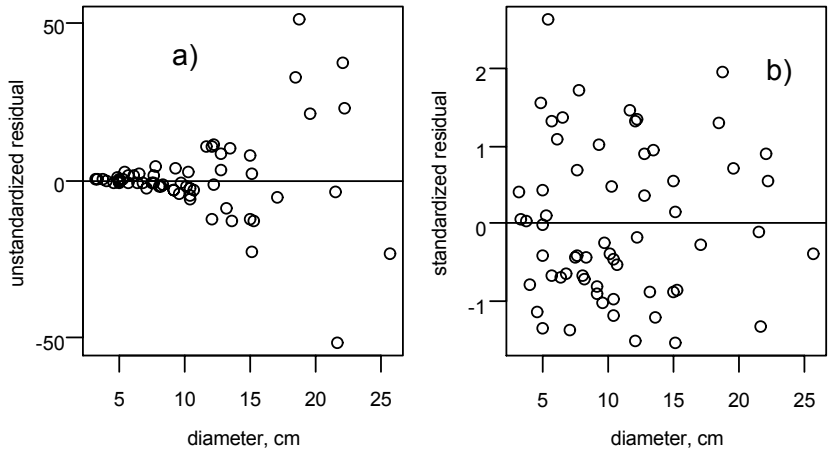

Figure 6.1 Unstandardized (a) and standardized (b) residuals of the weighted non-linear least squares fit. The standardized residuals were obtained by dividing the unstandardized residuals by the square root of the estimated variance function.

Example 6.2 Estimating plot-specific tally tree regression using a mixed model.

It can be clearly seen from Figure 6.2, in which the observed volumes for trees on the same sample plot are connected by dashed lines, that the residuals of the volume model are correlated for a single plot. Thus we assume a non-linear mixed model

$$y_{ki} = \exp(\alpha + a_k + b \ln x_{ki}) + e_{ki},$$

where y_{ki}, x_{ki} and e_{ki} are the volume, diameter and residual of tree i on plot k. Now $i = 1,\ldots,N_k$, where N_k is the total number of trees on plot k. The same model as in example 6.1 was assumed for the residual variance, the only difference being that parameter a was now defined as $\alpha + a_k$, i.e. it was assumed to be specific for each plot with the variance var(a_k) and expectation α. Fitting the model by means of restricted maximum likelihood gave the following parameter estimates:

$$\alpha = -2.001,\ b = 2.473,\ \mathrm{var}(a_k) = 0.1316^2,\ \sigma^2 = 0.006995^2 \text{ and } \delta = 2.548.$$

Predictions of a_k (McCulloch and Searle 2001, p. 247-262; Pinheiro and Bates 2000, p. 71) were calculated for each plot to obtain plot-specific volume models. The population curve and plot-specific curves for each plot are shown in Figure 6.2. Using the plot-specific curves, the RMSE of the volume prediction was 13.67dm^3, which shows a considerable reduction as compared with the model without random effects. This can be interpreted in two ways. From a statistical point of view, random effects can either be said to be part of the error term or can be considered to be random parameters. In the first interpretation, the original errors in the model with random effects are decomposed into two components, plot-level errors and tree-level errors, while in the second, the decrease in RMSE is obtained by adding a large number of additional parameters to the model.

In some cases it is desirable to use linear mixed models rather than non-linear ones. To demonstrate the difference between the two, the model

$$\ln y_{ki} = \alpha + a_k + b \ln x_{ki} + e_{ki}$$

was fitted to the data. Note that the first two terms of the model are obtained by taking a logarithm of the non-linear model without the error term. The difference between the non-linear and linearized models lies in the assumptions regarding the error term (see McCulloch and Searle 2001, p. 78). Fitting of the linearized model to the data gave the parameter estimates

$$\alpha = -2.002,\ b = 2.472,\ \mathrm{var}(a_k) = 0.1341^2 \text{ and } \mathrm{var}(e_{ki}) = 0.06067^2.$$

The estimates for the fixed parameters are almost equal to those of the non-linear model, the differences between the random parameters resulting from different assumptions concerning the error variance.

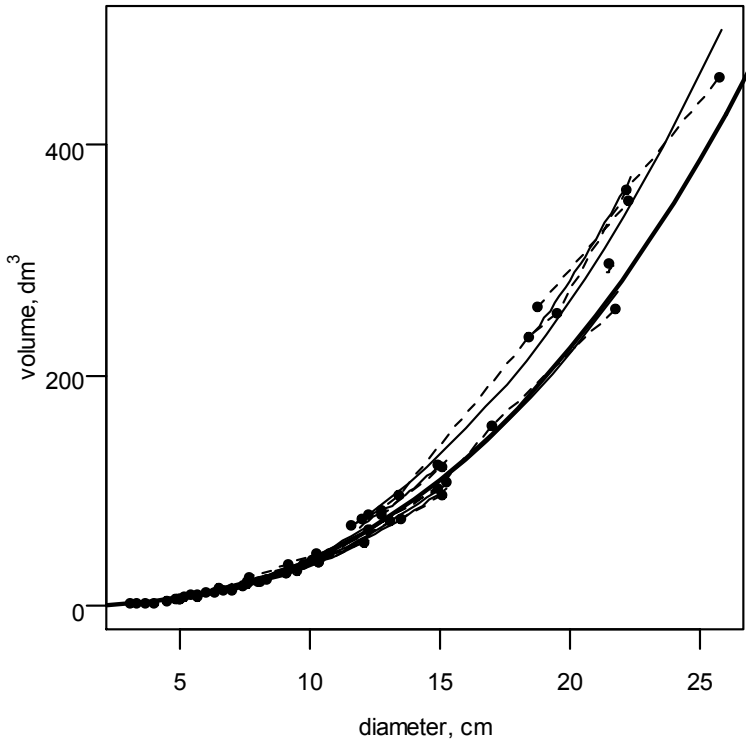

Figure 6.2 Observed volumes plotted against diameter. Observations from the same plot are connected by dashed lines. The thick solid line shows the population curve and the thin solid lines the plot-specific curves. Note that each plot-specific curve is plotted within the range of sample tree diameters of the plot and most of the 16 plot-specific curves are for diameters less than 15 cm.

6.2. GENERALIZING SAMPLE TREE INFORMATION IN A SMALL SUBPOPULATION

We may need to estimate $E(y|\mathbf{x})$ in a small population or subpopulation where not many sample trees have been measured. The subpopulations may be stands or small administrative areas, for instance. If we are using sample tree information collected only from the target subpopulation, the estimation errors in $\hat{E}(y|\mathbf{x})$ may be large. We may thus try to use some prior information to obtain a smaller MSE for $\hat{E}(y|\mathbf{x})$, even though the estimate for $\hat{E}(y|\mathbf{x})$ may be biased for the given subpopulation. Two such methods will be shortly discussed in the following, mixed estimation and mixed models.

6.2.1 Mixed estimation

Mixed estimation can be applied if there are two datasets available: a dataset sampled from the current target population and an auxiliary dataset from a population which we anticipate to be quite similar to the target population. The two are then combined, with less weight attached to the observations in the auxiliary dataset than to those in the target population. Korhonen (1993) used mixed estimation to calibrate the data of the 8th Finnish National Forest Inventory using data from the previous inventory.

More specifically: let $\mathbf{y}_1$ and $\mathbf{y}_2$ be the vectors of the dependent variables for the sample as obtained from the target population and from the auxiliary population, respectively. Using the example above, these could be the values of the sample tree volume equation $V(D,H)$, possibly divided by a function $s(D)$ used to stabilize the error variance. Let us assume a linear model $\mathbf{y}_1 = \mathbf{X}_1\mathbf{b} + \mathbf{e}$ and let $\mathbf{X}_1$ and $\mathbf{X}_2$ be the model matrices from the two samples, respectively. In mixed estimation $\mathbf{b}$ is estimated by $\hat{\mathbf{b}} = \left(\mathbf{X}_1'\mathbf{X}_1 + \lambda\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\left(\mathbf{X}_1'\mathbf{y}_1 + \lambda\mathbf{X}_2'\mathbf{y}_2\right)$ instead of the ordinary least squares estimate $\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{y}_1$. Thus a weighted least squares regression is applied in which the weights on the observations of the target population are 1 and those on the observations of the auxiliary population are λ. If the two populations do not deviate much, the resulting estimate will have a smaller MSE even if it is biased. For more information on mixed estimation, see Theil and Goldberger (1961) and Toutenburg (1982). Use of the mixed estimation approach with the non-linear volume model of Example 6.1 is illustrated in Example 6.3.

Ridge regression is another biased estimation method that can provide parameter estimates having a smaller MSE error than OLS, especially in small datasets. In this method the parameter estimates are shrunken towards zero. Both mixed estimation and ridge regression are formally similar to the prediction of random parameters using mixed models. Mixed models have a more natural theoretical basis, however.

6.2.2 Applying mixed models

If we are making an inventory in several subpopulations or classes which can be considered a priori to be similar, we can employ techniques which 'borrow strength' from other subpopulations. The classes can be sample plots or stands, for instance. The fitting of a mixed model was discussed in section 6.1. We now show how an estimated mixed model can be employed to predict tally tree regressions for small subpopulations outside the estimation data.

Mixed model prediction is based on linear prediction theory, which we will first explain briefly. Assume that we have a vector of random variables, **h**, which can be divided into two parts

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix},$$

where $\mathbf{h}_1$ and $\mathbf{h}_2$ are random vectors of length 1 or more. It is assumed that $E(\mathbf{h}_1)=\boldsymbol{\mu}_1$, $E(\mathbf{h}_2)=\boldsymbol{\mu}_2$, $\text{var}(\mathbf{h}_1)=\mathbf{V}_1$, $\text{var}(\mathbf{h}_2)=\mathbf{V}_2$, and $\text{cov}(\mathbf{h}_1,\mathbf{h}_2\text{'})=\mathbf{V}_{12}$. Using the notation of McCulloch and Searle (2001, p. 247), this can be written as

$$\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_{12} \\ \mathbf{V}_{12}\text{'} & \mathbf{V}_2 \end{bmatrix} \right). \tag{6.1}$$

Assume that we have observed the random vector $\mathbf{h}_2$ and want to predict vector $\mathbf{h}_1$. The Best Linear Predictor of $\mathbf{h}_1$ is

$$BLP(\mathbf{h}_1) = \hat{\mathbf{h}}_1 = \boldsymbol{\mu}_1 + \mathbf{V}_{12}\mathbf{V}_2^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \tag{6.2}$$

with a prediction variance of

$$\text{var}(\hat{\mathbf{h}}_1 - \mathbf{h}_1) = \mathbf{V}_1 - \mathbf{V}_{12}\mathbf{V}_2^{-1}\mathbf{V}_{12}\text{'} \tag{6.3}$$

(McCulloch and Searle 2001, p. 250). Thus, if the expectations and variance-covariance matrices of two random vectors are known and either one of them is observed, the other one can be predicted, and the variance of the prediction error can be calculated using Equation (6.3). If $\mathbf{h}$ follows the multinormal distribution, BLP will also be the Best Predictor. If the matrices $\mathbf{V}_1$, $\mathbf{V}_2$ and $\mathbf{V}_{12}$ and the vector $\boldsymbol{\mu}_2$ are replaced in the calculations by their estimates, the resulting predictor is the Estimated Best Linear Unbiased Predictor (EBLUP). Henderson mixed model equations lead to equivalent equations that do not require the inverse of $\mathbf{V}_2$ and are thus more convenient when the number of sample trees in the subpopulation is large (see Searle et al. 1991, Lappi 1991).

Assume that the sample tree variables in subpopulation k are written in vector $\mathbf{y}_k$, and assume that they follow a mixed model

$$\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k, \beta) + \mathbf{Z}\mathbf{b}_k + \mathbf{e}_k, \tag{6.4}$$

where $\mathbf{f}(\mathbf{x}_k, \beta)$ is the fixed part of the model, giving $E(\mathbf{y}|\mathbf{x}_k)$, $\mathbf{b}_k$ includes the random parameters for subpopulation k, $\mathbf{Z}$ is the design matrix corresponding to the random coefficients and $\mathbf{e}_{ki}$ is the vector of residuals. The design matrix has a row for each sample tree, which includes those observed predictors which have a random coefficient. Linearity in the random part is required for simplicity. For prediction of the random parameters of non-linear mixed models, see Pinheiro and Bates (2000) and Fang et al. (2001).

Let us note var(**b**)=**D** and var(**e**)=**R.** Estimates for these matrices are available from the estimation stage of the model. If the errors are homogeneous and uncorrelated, **R**=σ2**I**. Let us define $\mathbf{h}_1=\mathbf{b}_k$, $\mathbf{h}_2=\mathbf{y}_k$ in equation (6.1). Using model (6.4) we obtain $\mathbf{V}_1=\mathbf{D}$, $\mathbf{V}_2=\mathbf{ZDZ'}+\mathbf{R}$ and $\mathbf{V}_{12}=\mathbf{DZ'}$. Now we can use equation (6.2) to predict the random parameters in subpopulation k, denoted by $\hat{\mathbf{b}}_k$, and equation (6.3) to calculate its prediction error, $\mathrm{var}\left(\hat{\mathbf{b}}_k-\mathbf{b}_k\right)$.

According to model (6.4), the prediction regarding tree i in subpopulation k is

$$\hat{y}_{ki}=E\left(y\left|\mathbf{x}_{ki},\mathbf{y}_k\right.\right)=f\left(\mathbf{x}_{ki},\beta\right)+\mathbf{z}\hat{\mathbf{b}}_k\ .$$

The variance in the predictions is needed for calculating their confidence intervals, for example, and for correcting for bias if y_{ki} is not on an arithmetic scale. Ignoring the estimation error in the fixed parameters, the prediction variance is

$$\mathrm{var}\left(\hat{y}_{ki}-y_{ki}\right)=\mathbf{z}\,\mathrm{var}\left(\hat{\mathbf{b}}_k-\mathbf{b}_k\right)\mathbf{z'}+\mathrm{var}\left(e_{ki}\right),$$

where **z** is the row of **Z** that corresponds to the observation $\mathbf{x}_{ki}$.

The above approach can be used when the sample tree variables in the modelling data coincide with those in the target population data. Using the volume example, predicted sample tree volumes based on diameter and height can be used if the volumes in the modelling data are based on same volume model. This is not the case, however, if the volumes in the modelling data are true volumes based on stem analysis and the sample tree volumes are predictions based on diameter and height. A multivariate approach for predicting sample tree volumes using measured sample tree heights, developed by Lappi (1991), is presented in Example 6.4.

Example 6.3 Generalizing sample tree information to a small subpopulation

Assume that we want to generalize the volumes for tally trees on a plot that is not included in the modelling data of the previous examples. Assume that two sample trees were measured on the plot (plot 51 in Table 1). This example demonstrates the use of both mixed estimation and mixed model prediction for generalizing sample tree information.

Mixed estimation

In order to homogenize the residual errors, the model of Example 6.1 was transformed to

$$\frac{y_i}{s(x_i)} = \frac{1}{s(x_i)}\exp(a + b\ln x_i) + e_i ,$$

where, based on the variance function of Example 6.1, $s(x_i) = x_i^{2.68}$.

The dataset used in the previous examples was taken here as the auxiliary dataset for mixed estimation. Since the auxiliary data include 61 trees and the plot data only 2, uniform weighting of all the observations in the combined dataset would have given too much weight to the auxiliary data. Defining the weights as $\lambda=1$ in the plot data and $\lambda=2/61=0.0328$ in the auxiliary data gives half of the weight to the plot data and half to the auxiliary data. Fitting of the model to the combined data using weighted least squares gave the parameter estimates

$$a = -2.400 \text{ and } b = 2.631.$$

The resulting curve is shown by the dashed line in Figure 6.3. One can see that it is more accurate than the population curve of Example 6.1, but clearly gives excessively large volumes for the sample trees. The RMSE of the prediction was 12.51 dm^3 and the bias (predicted-observed) 8.80 dm^3.

Another approach to mixed estimation would have been to generate an artificial auxiliary dataset using an existing volume model (Pekkonen 1982). In this case, for example, one could have used the model of Example 6.1 to generate artificial trees with diameters varying systematically within the diameter range of the tally trees. The mixed estimation model would then have been fitted to the dataset consisting of the observed sample trees and the artificial sample trees and the weighting of the auxiliary data could have been controlled by the number of artificial trees.

Mixed model prediction

In this example the random parameters of the linearized volume model of Example 6.2 were predicted for the plot using the two measured sample trees of plot 51 (Table 6.1). Since the only random parameter in our example is the constant term, the average of the residuals includes all the information that is needed for predicting the random effect and equations (6.1 and 6.2) can be simplified. Let us first calculate the average of the observed logarithmic volumes as $\bar{y}_k = \frac{1.21+2.87}{2} = 2.04$. The average of the expected logarithmic volumes is obtained using fixed parameters of the model: $\mu = \frac{1.36+3.17}{2} = 2.27$. Equation (6.1) now takes the form

$$\begin{bmatrix} a_k \\ \bar{y}_k \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ \mu \end{bmatrix}, \begin{bmatrix} \operatorname{var}(a_k) & \operatorname{var}(a_k) \\ \operatorname{var}(a_k) & \operatorname{var}(a_k) + \operatorname{var}(e_{ki})/n_k \end{bmatrix} \right),$$

where $n=2$, $\operatorname{var}(a_k) = 0.134^2$ and $\operatorname{var}(e_{ki}) = 0.0607^2$. The Best Linear Unbiased Predictor of a_k is calculated to be

$$\hat{a}_k = \frac{n_k \operatorname{var}(a_k)}{\operatorname{var}(e_{ki}) + n_k \operatorname{var}(a_k)} (\bar{y}_k' - \mu) = \frac{2 \cdot 0.134^2}{0.0607^2 + 2 \cdot 0.134^2} (2.04 - 2.27) = -0.206 .$$

Thus, according to the model (see example 6.2), the expected logarithmic volume on the plot is

$$E(\ln y_{ki} \mid \bar{y}_k) = -2.002 - 0.206 + 2.47 \ln x_{ki} = -2.208 + 2.47 \ln x_{ki} .$$

Before applying the exponential transformation to obtain the volumes on an arithmetic scale, half of the error variance needs to be added to the logarithmic prediction. The error variance consists of the residual error $\operatorname{var}(e_{ki}) = 0.0607^2$ and the prediction error of the plot effect,

$$\operatorname{var}(\hat{a}_k - a_k) = \operatorname{var}(a_k) - \frac{n_k \left[\operatorname{var}(a_k) \right]^2}{\operatorname{var}(e_{ki}) + n_k \operatorname{var}(a_k)} = 0.134^2 - \frac{2 \cdot 0.134^4}{0.0607^2 + 2 \cdot 0.134^2}$$
$$= 0.00167.$$

Thus, the volume expectation is

$$E(y_{ki} \mid \bar{y}_k) = \exp\left[-2.208 + \frac{0.00167 + 0.00368}{2} + 2.47 \ln x_{ki} \right]$$
$$= \exp(-2.202 + 2.47 \ln x_{ki}),$$

which is shown with a solid line in Figure 6.3. The predictions are more accurate than in the mixed estimation approach (RMSE 4.43), but they are still slightly biased for plot 51, the observed bias being 3.01 dm^3.

Both the mixed estimation and mixed model prediction approaches lead to a plot-specific volume model that is obtained from the population curve by shrinking it towards the observed volumes of the sample trees. In mixed estimation the degree of shrinkage depends on the *ad hoc* weights assigned to the auxiliary data, while in the mixed model approach it is defined by the estimated within-plot and between-plot variances in the mixed model.

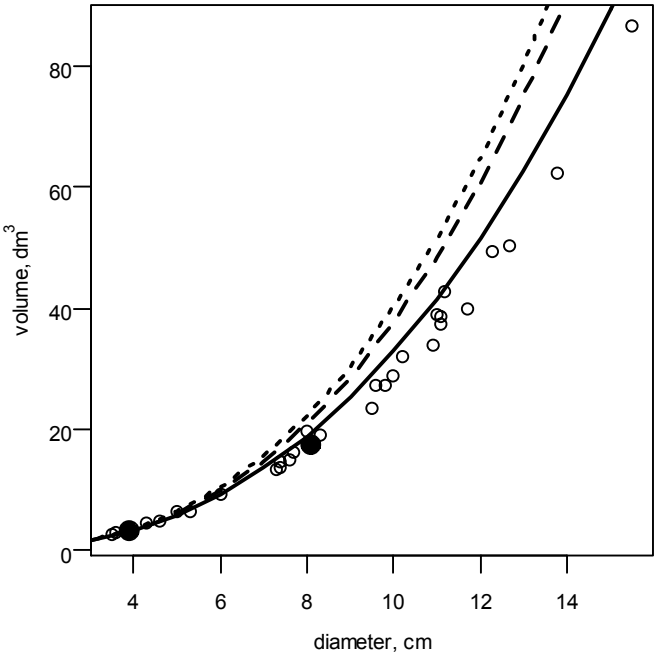

Figure 6.3 The dataset and fitted curves of Examples 6.1 and 6.3. The black circles show the true volumes of the two sample trees and the open circles those of the tally trees. The dotted line shows the expected curve of Example 6.1, the dashed line the curve obtained with mixed estimation and the solid line that obtained with mixed model prediction.

Example 6.4 The multivariate case

Lappi (1991) constructed the following multivariate model for the logarithmic height and logarithmic volume of tree i in stand k from stem analysis data (Laasasenaho 1982):

$$\ln H_{ki} = 3.410 - 18.58\frac{1}{D_{ki}} + a_{0k} - a_{1k}\frac{1}{D_{ki}} + e_{ki} \text{ and}$$

$$\ln V_{ki} = 2.704 - 48.93\frac{1}{D_{ki}} + 1.387\ln D_{ki} + c_{0k} - c_{1k}\frac{1}{D_{ki}} + u_{ki},$$

where D_{ki} is DBH + 7 cm, parameters a_{0k}, a_{1k}, c_{0k} and c_{1k} are stand-specific random parameters and e_{ki} and u_{ki} are residuals with estimated variances of $\text{var}(e_{ki}) = 0.01113$, $\text{var}(u_{ki}) = 0.01540$ and covariance $\text{cov}(e_{ki}, u_{ki}) = 0.01040$. Let us write the random

parameters as vectors $\mathbf{a}_k = (a_{0k} \quad a_{1k})$' $\mathbf{c}_k = (c_{0k} \quad c_{1k})$' and define $\mathbf{b}_k = (\mathbf{a}_k$' $\mathbf{c}_k$')'. The estimated dispersion matrix of $\mathbf{b}_k$ is (Lappi 1991)

$$\operatorname{var}(\mathbf{b}_k) = \mathbf{D} = \left[\begin{array}{cc|cc} 0.04739 & -0.3887 & 0.05082 & -0.4772 \\ -0.3887 & 20.64 & -0.6036 & 24.88 \\ \hline 0.05082 & -0.6036 & 0.05988 & -0.7876 \\ -0.4772 & 24.88 & -0.7876 & 31.11 \end{array}\right]$$

$$= \left[\begin{array}{c|c} \mathbf{D}_1 & \mathbf{D}_{12}\text{'} \\ \hline \mathbf{D}_{12} & \mathbf{D}_2 \end{array}\right] = [\mathbf{C} \quad \mathbf{H}].$$

The last two parts define a partition of matrix **D** that is needed in the following calculations. The measured height of a sample tree will be used below to predict the random parameters of the volume function. Assume that two sample trees of diameters 20 and 30 cm and heights 20 and 26 m have been measured. The measured heights follow the model

$$\mathbf{y}_k = \boldsymbol{\mu} + \mathbf{Z}\mathbf{a}_k + \mathbf{e}_k ,$$

where vector $\mathbf{y}_k$ includes the measured logarithmic heights, $\mathbf{y}_k = \begin{bmatrix} \ln 20 \\ \ln 26 \end{bmatrix} = \begin{bmatrix} 3.00 \\ 3.26 \end{bmatrix}$, and $\boldsymbol{\mu}$ their expectations, which are obtained using the first two terms of the height model as $\boldsymbol{\mu} = \begin{bmatrix} 2.72 \\ 2.91 \end{bmatrix}$. Matrix **Z** is the design matrix of the random part, i.e., $\mathbf{Z} = \begin{bmatrix} 1 & 1/(20+7) \\ 1 & 1/(30+7) \end{bmatrix}$, and $\mathbf{a}_k$ and $\mathbf{e}_k$ are unknown vectors of random parameters and random residuals with variances $\operatorname{var}(\mathbf{a}_k) = \mathbf{D}_1$ and $\operatorname{var}(\mathbf{e}_k) = \mathbf{R} = 0.01113 \cdot \mathbf{I}$. Using the height and volume models, equation (6.1) can be written as

$$\begin{bmatrix} \mathbf{b}_k \\ \mathbf{y}_k \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \boldsymbol{\mu} \end{bmatrix}, \begin{bmatrix} \mathbf{D} & \mathbf{C}\mathbf{Z}\text{'} \\ \mathbf{Z}\mathbf{C}\text{'} & \mathbf{Z}\mathbf{D}_1\mathbf{Z}\text{'} + \mathbf{R} \end{bmatrix} \right)$$

and the BLUP of $\mathbf{b}_k$ is (Equation 6.2)

$$\hat{\mathbf{b}}_k = \mathbf{CZ'}\left(\mathbf{ZD}_1\mathbf{Z'}+\mathbf{R}\right)^{-1}\left(\mathbf{y}_k-\boldsymbol{\mu}\right) = \begin{pmatrix} 0.244 \\ 0.985 \\ 0.230 \\ 1.131 \end{pmatrix},$$

i.e. the predicted random parameters are $a_{0k} = 0.244$, $a_{1k} = 0.985$, $c_{0k} = 0.230$ and $c_{1k} = 1.131$. The predicted logarithmic heights and volumes are obtained by writing these estimates into the height and volume models.

In order to arrive at unbiased predictions of volumes and heights, half of the prediction variance was added to the predicted logarithmic heights and volumes before applying the exponential transformation. The prediction variance of random parameters was first calculated to be

$$\begin{aligned} \operatorname{var}\left(\hat{\mathbf{b}}_k - \mathbf{b}_k\right) &= \mathbf{D} - \mathbf{ZC'}\left(\mathbf{ZD}_1'\mathbf{Z} + \mathbf{R}\right)^{-1}\mathbf{CZ'} \\ &= \begin{pmatrix} 0.0212 & -0.536 & 0.0266 & -0.648 \\ -0.536 & 17.6 & -0.716 & 21.3 \\ 0.0266 & -0.716 & 0.0372 & -0.918 \\ -0.648 & 21.3 & -0.918 & 26.8 \end{pmatrix}. \end{aligned}$$

Ignoring the estimation errors in the fixed parameters, the prediction variances of the predicted logarithmic heights were then obtained from the diagonal of

$$\operatorname{var}\left(\hat{\mathbf{y}}_k\,* - \mathbf{y}_k\,*\right) = \mathbf{Z}*\operatorname{var}\left(\hat{\mathbf{a}}_k - \mathbf{a}_k\right)\mathbf{Z}*' + 0.01113\mathbf{I},$$

where $\mathbf{y}_k$* denotes the heights of the tally trees, **Z*** the design matrix of tally trees and $\operatorname{var}\left(\hat{\mathbf{a}}_k - \mathbf{a}_k\right)$ includes the first two rows and columns of $\operatorname{var}\left(\hat{\mathbf{b}}_k - \mathbf{b}_k\right)$ (see the definition of $\mathbf{b}_k$). The height and volume models corrected for population level and local bias are shown in Figure 6.4.

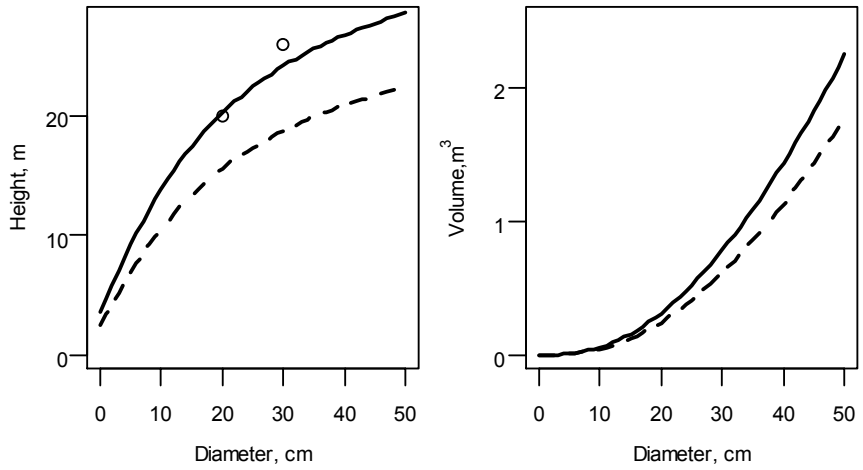

Figure 6.4 Predicted height and volume models when random parameters are 0 (dashed lines) and are predicted using the two observed heights shown in the plot on the left.

6.3 A CLOSER LOOK AT THE THREE-LEVEL MODEL STRUCTURE

Let y denote a variable for which we do not have any measurements in the inventory data (e.g. stem volume), let $\mathbf{x}$ denote tally tree measurements (e.g. dbh), and let $\mathbf{z}$ denote sample tree variables for which we have some measurements (e.g. tree height). The statistical analysis is straightforward if we assume that both $\mathbf{x}$ and $\mathbf{z}$ are random vectors. We then assume that we know the conditional expectation $E\left(y|\mathbf{x},\mathbf{z}\right)$ (called here the sample tree regression) and the conditional variance $\mathrm{var}\left(y|\mathbf{x},\mathbf{z}\right)$, which have been estimated from previous research data. The sample tree regression can be applied using sample tree measurements (note that $\mathrm{var}\left(y|\mathbf{x},\mathbf{z}\right)=\mathrm{var}\left(y-E\left(y|\mathbf{x},\mathbf{z}\right)|\mathbf{x},\mathbf{z}\right)$). We would like to estimate the tally tree regression $E\left(y|\mathbf{x}\right)$ and its residual variance $\mathrm{var}\left(y|\mathbf{x}\right)$ using a random tree sample from the target population.

The conditional expectation $E\left(y|\mathbf{x}\right)$ can be presented as (see Rao 1973, p. 97)

$$E\left(y|\mathbf{x}\right)=E_{\mathbf{z}|\mathbf{x}}E\left(y|\mathbf{x},\mathbf{z}\right) \tag{6.5}$$

where $E_{\mathbf{z}|\mathbf{x}}$ denotes the expectation over the distribution of $\mathbf{z}$ for a given value of $\mathbf{x}$.

There are two approaches for estimating the tally tree regression $E(y|\mathbf{x})$ using (6.5). First, if we have a random sample from the population, the tally tree regression $E(y|\mathbf{x})$ can be estimated using the sample tree regression function $E(y|\mathbf{x},\mathbf{z})$ as the dependent variable which is regressed on $\mathbf{x}$, as in Examples 6.1 and 6.2. The multivariate approach presented in Example 6.4 can also be shown to be equivalent to this approach.

A second possibility is to estimate the conditional distribution of **z** for a given value of **x** and then compute the expected value for $E(y|\mathbf{x},\mathbf{z})$ with respect to this conditional distribution. If $E(y|\mathbf{x},\mathbf{z})$ is not linear with respect to **z**, a numerical integration or approximation method is needed to compute the expected value.

The prediction variance of *y* can be expressed as (see Rao 1973):

$$\operatorname{var}(y|\mathbf{x}) = E_{\mathbf{z}|\mathbf{x}} \operatorname{var}(y|\mathbf{x},\mathbf{z}) + \operatorname{var}_{\mathbf{z}|\mathbf{x}} E(y|\mathbf{x},\mathbf{z}). \tag{6.6}$$

When $E(y|\mathbf{x})$ is estimated by regressing observed values of $E(y|\mathbf{x},\mathbf{z})$ on $\mathbf{x}$, the estimator $\hat{E}(y|\mathbf{x})$ also entails some estimation error. When *y* is predicted for tally trees using $\hat{E}(y|\mathbf{x})$, i.e. $\hat{y} = \hat{E}(y|\mathbf{x})$, the expected squared error consists of both the estimation error of the regression model ($\hat{E}(y|\mathbf{x}) - E(y|\mathbf{x})$) and the residual variance $\operatorname{var}(y|\mathbf{x})$ of the true model $E(y|\mathbf{x})$. Thus, for a given $\hat{E}(y|\mathbf{x})$:

$$E\left((\hat{y}-y)^2|\mathbf{x}\right) = \left(\hat{E}(y|\mathbf{x}) - E(y|\mathbf{x})\right)^2 + \operatorname{var}(y|\mathbf{x}). \tag{6.7}$$

The estimation error $\hat{E}(y|\mathbf{x}) - E(y|\mathbf{x})$ consists of both the random error of the estimation procedure and possible model mis-specification error, i.e. a bias. If the bias component can be assumed to be small, the expected squared error over repeated estimations of $\hat{E}(y|\mathbf{x})$ will be

$$E\left((\hat{y}-y)^2|\mathbf{x}\right) = \operatorname{var}\left(\hat{E}(y|\mathbf{x})\right) + E_{\mathbf{z}|\mathbf{x}} \operatorname{var}(y|\mathbf{x},\mathbf{z}) + \operatorname{var}_{\mathbf{z}|\mathbf{x}} E(y|\mathbf{x},\mathbf{z}). \tag{6.8}$$

The first component of (6.8) results from the estimation errors in the model and the last two components are the residual errors of sample tree regression and tally tree regression, respectively**.**

As an example, let us assume that dbh, denoted as *D,* is the tally tree measurement, and dbh and height, denoted as *H*, are sample tree measurements, and we want to predict the volumes of the tally trees, denoted as *V*. The known sample tree volume equation is denoted by $V(D,H)$ and the unknown tally tree volume equation by $V(D)$, that is $V(D,H)=E(V|D,H)$, and $V(D)=E(V|D)$. The relation between *V* and *D* is dependent on the history of the stands, so that we cannot assume that there exists a universal relationship $V(D)$ which is constant from inventory to inventory. Using (6.5), we obtain

$$V(D)=E_{H|D}V(D,H).$$

Adopting the first approach, the volume of the tally trees can be predicted by constructing a regression function for $V(D,H)$ with respect to *D*. Taking the estimation error of $V(D)$ into account as well, the expected squared error is (Equation 6.8):

$$E\left(\hat{V}(D)-V\right)^2=\text{var}\left(\hat{V}(D)-V(D)\right)+E_{H|D}\,\text{var}\left(V|D,H\right)+\text{var}_{H|D}\,E\left(V|D,H\right).$$

(6.9)

The conditional variances may be easier to understand if they are presented in an equivalent form showing that they are variances of residual errors:

$$\text{var}\left(V|D,H\right)=\text{var}\left(V-V(D,H)\right) \quad \text{and}$$

$$\text{var}_{H|D}\,E\left(V|D,H\right)=\text{var}\left(V(D,H)-V(D)\right)$$

The last two terms in (6.9) are now the residual variance of the sample tree volume equation (estimated from true volumes in original research data) and the variance of the sample tree volume function around its expected value, i.e. the residual variance of the tally tree volume equation. Referring to Example 6.1, the first term is the estimation variance of the model, the second term is the residual variance of the volume equation of Laasasenaho (1982) and the third term is the residual variance of the estimated model.

Adopting the second approach, we first estimate the conditional distribution of *H* for a given *D* and then compute the expected value of $V(D,H)$ with respect to this distribution. The most important property of the conditional distribution of *H* for a given *D* is the expected value $H(D)$, i.e. the height/dbh curve. If the distribution

of H for a given D is approximated by its expected value, i.e. the whole probability mass is shifted to the expected value, we obtain the common simple approach in which $V\left(D,H\left(D\right)\right)$ is used to predict tally tree volumes (Clutter et al. 1983, West 2004). This is justified if $V\left(D,H\right)$ is linear with respect to H, that is $V\left(D,H\right)=g\left(D\right)H$ for a known function of g. In this case $E\left(V\left(D,H\right)\middle|D\right)=g\left(D\right)H\left(D\right)$ (see Rao 1973, p. 97). But because the volume is slightly non-linear with respect to height for a given dbh, the use of $V\left(D,H\left(D\right)\right)$ entails a certain bias. If the error variance is to be computed using only the error variance of $V\left(D,H\right)$, the third term in (6.9) will be ignored. For an error analysis of this approach, see Gregoire and Williams (1992).

If the error variance of the height equation $H\left(D\right)$ is also estimated, a simple approximation for the distribution of H for a given D is a two-point distribution which places half of the probability mass on the expected value plus standard deviation and half on the expected value minus the sd. Using this two point distribution, we arrive at the approximations:

$$E\left(V\right)\approx 0.5\left(E\left(V\middle|D=d,H=\mu-\sigma\right)+E\left(V\middle|D=d,H=\mu+\sigma\right)\right) \qquad (6.10)$$

$$\begin{aligned}\operatorname{var}\left(V\right)\approx{}& 0.5\left(\operatorname{var}\left(V\middle|D=d,H=\mu-\sigma\right)+\operatorname{var}\left(V\middle|D=d,H=\mu+\sigma\right)\right)+\\ & 0.5\left(\left(E\left(V\middle|D=d,H=\mu-\sigma\right)-E\left(V\right)\right)^2+\left(E\left(V\middle|D=d,H=\mu+\sigma\right)-E\left(V\right)\right)^2\right),\end{aligned} \qquad (6.11)$$

where d is the diameter of the tally tree, μ its expected height and σ the standard deviation of the height prediction (Example 6.5).

The above equations can be applied to individual trees if the trees in the inventory data are assumed to be independent. The assumption of a model with plot-level random effects, for example, implies that the trees on the same plot are correlated, and the conditional expectations must be computed by taking into account the tally tree and sample tree variables for all the trees on the same plot.

Example 6.5 Use of the distribution of H|D to estimate tally tree volumes.

Recalling the calibrated height model of Example 6.4, let us now assume that we are predicting the volumes of tally trees of diameters 10, 20 and 30 cm. The design matrix of the last equation of Example (6.4) is

$$\mathbf{Z}^* = \begin{pmatrix} 1 & 1/(10+7) \\ 1 & 1/(20+7) \\ 1 & 1/(30+7) \end{pmatrix}$$

and the variance-covariance matrix of the prediction error becomes

$$\mathrm{var}\left(\ln \hat{\mathbf{h}}_k - \ln \mathbf{h}_k\right) = \begin{pmatrix} 0.130 & 0.00820 & 0.00319 \\ 0.00820 & 0.117 & 0.00449 \\ 0.00319 & 0.00449 & 0.116 \end{pmatrix}.$$

The expectations for the logarithmic heights of the tally trees are shown in the second column of Table 6.2 and the standard deviation of their prediction errors, obtained from the diagonal of $\mathrm{var}\left(\ln \hat{\mathbf{h}}_k - \ln \mathbf{h}_k\right)$, in the third column. The volumes of the tally trees were first calculated by the traditional approach, using the expected height (H_1) to predict the volumes (V_1) and applying bias correction in the prediction of the heights (Table 6.3). Ignoring the height prediction error, the standard error in prediction (sd(V_1)) was calculated to be $\mathrm{sd}\left(V \middle| D,H\right) \approx 0.075E\left(V \middle| D,H\right)$ (Laasasenaho 1982). In the second approach, two predicted heights were calculated for each tally tree: one obtained by subtracting the standard deviation of the prediction from (H_2) and the other by adding it to (H_3), the prediction of the expected log height before exponential transformation to calculate the volumes of the tally trees (V_2 and V_3). The final volume estimates, V_4, were obtained as averages of V_2 and V_3 (Equation 6.10). The prediction variance (sd(V_4), Equation 9.11), takes into account both the height prediction error and the volume function prediction error .

There are only slight differences in the point estimates of volume, but considerable differences in the prediction errors. The merits of the second approach over the first are that a realistic estimate of the prediction variance is obtained and that bias correction based on Taylor series approximation and normality of the errors in the prediction of log height is not needed.

Table 6.2 Predicted heights and volumes of the three tally trees and their prediction errors.

D	ln$H$	sd(ln$H$)	H_1	H_2	H_3
10	2.619	0.1739	13.93	11.53	16.33
20	3.002	0.1266	20.30	17.69	22.92
30	3.179	0.1274	24.21	21.14	27.27

Table 6.3 Predicted volumes of the three tally trees and their prediction errors.

D	V_1	V_2	V_3	V_4	sd(V_1)	sd(V_4)
10	55.56	47.09	64.06	55.57	4.167	9.475
20	305.7	269.3	342.2	305.8	22.93	43.13
30	788.9	695.3	882.9	788.9	59.17	111.0

REFERENCES

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., Bailey, R.L. 193. Timber management. A quantitative approach. Wiley, New York. 333 p.

Davidian, M. and Giltinan, D.M. 1995. Nonlinear models for repeated measurement data. Monographs on statistics and applied probability 62. Chapman & Hall/CRC, London. 359 p.

Diggle, P.J., P. Heagerty, K.-Y. Liang, and S.L. Zeber. 2002. Analysis of longitudinal data. Second edition. Oxford University Press. 379 pp.

Fang, Z, Bailey, R.L. and Shiver, B.D. 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. Forest Science 47: 550-562.

Gregoire, T.G. and Williams, M. 1992. Identifying and evaluating the components of non-measurement error in the application of standard volume equations. The Statistician 41: 509-518.

Gregoire, T.G. and Schabenberger, O. 1996. A non-linear mixed-effects model to predict cumulative bole volume of standing trees. Journal of Applied Statistics 23: 257-271.

Gregoire, T.G., Schabenberger, O. and Barrett, J.P. 1995. Linear modeling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Canadian Journal of Forest Reserach 25: 137-156.

Gustavsen H.G. 1998. Volymtillväxten och övre höjdens utveckling i talldominerande bestånd i Finland – en utvärdering av några modellers validitet i nuvarande skogar. The Finnish Forest Research Institute. Research Notes 707.

Hastie, T, Tibshirani R. and Friedman, J. 2001. The elements of statistical learning: Data Mining, Inference and prediction. Springer, New York. 533pp.

Kangas, A. 1997. On the prediction bias and variance in long-term growth projections. Forest Ecology and Managagement 96: 207-216.

Korhonen, K.T. and Kangas, A. 1997. Application of nearest-neighbour regression for generalizing sample tree information. Scandinavian Journal of Forest Research 12: 97-101.

Korhonen, K.T. 1993. Mixed estimation in calibration of volume functions of Scots pine. Silva Fennica 27: 269-276.

Laasasenaho, J. 1982. Taper curve and volume functions for pine, spruce and birch. Communicationes Instituti Forestalis Fenniae 108.

Lappi, J. 1991. Calibration of height and volume equations with random parameters. Forest Science 37: 781-801

Lappi, J. 1991. Calibration of height and volume equations with random parameters. Forest Science 37: 781-801.

Lappi, J. 1993. Metsäbiometrian menetelmiä. Silva Carelica 24. University of Joensuu, 182 p.

McCulloch, C. E. and Searle, S. R. 2001. Generalized, linear, and mixed models. John Wiley & Sons. 325 p.

Mehtätalo, L. 2004. A longitudinal height-diameter model for Norway spruce in Finland. Canadian Journal of Forest Research 34: 131-140.

Pekkonen, T. Leimikon puuston tilavuuden arviointi regressioennustinta käyttäen. [Regression estimators in predicting the volume of a stand marked for cutting. In Finnish]. Finnish Forest Research Institute. Research Notes 86.

Pinheiro, J.C. and Bates, D.M. 2000. Mixed-Effects Models in S and S-PLUS. Springer, New York. 528 p.

Rao, C.R. 1973. Linear statistical inference and its applications. Wiley, New York. 625 pp.

Searle, S.R., Casella, G. and McCulloch, C.E. 1991. Variance components. Wiley, New York. 501 p.

Theil, H. and Goldberger, A.S. 1961. On pure and mixed statistical estimation in economics. International Economical Review 2:65-78

Toutenburg, H. 1982. Prior information in linear models. Wiley. Chischester. 215 pp.

West, P.W. 2004. Tree and forest measurement. Springer, Berlin. 167 p.