Eur J Forest Res (2007) 126:573–585
DOI 10.1007/s10342-007-0179-1

ORIGINAL PAPER

A variance-covariance structure to take into account repeated measurements and heteroscedasticity in growth modeling

Mathieu Fortin · Gaétan Daigle · Chhun-Huor Ung · Jean Bégin · Louis Archambault

Received: 31 January 2006 / Accepted: 23 April 2007 / Published online: 31 May 2007

Abstract This study proposes a within-subject variance-covariance (VC) structure to take into account repeated measurements and heteroscedasticity in a context of growth modeling. The VC structure integrates a variance function and a continuous autoregressive covariance structure. It was tested on a nonlinear growth model parameterized with data from permanent sample plots. Using a stand-level approach, basal area growth was independently modeled for red spruce (*Picea rubens* Sarg.) and balsam fir [*Abies balsamea* (L.) Mill.] in mixed stands. For both species, the implementation of the VC structure significantly improved the maximum likelihood of the model. In both cases, it efficiently accounted for heteroscedasticity and autocorrelation, since the normalized residuals no longer exhibited departures from the assumptions of independent error terms with homogeneous variances. Moreover, compared with traditional nonlinear least squares (NLS) models, models parameterized with this VC structure may generate more accurate predictions when prior information is available. This case study demonstrates that the implementation of a VC structure may provide parameter estimates that are consistent with asymptotically unbiased variances in a context of nonlinear growth modeling using a stand-level approach. Since the variances are no longer biased, the hypothesis tests performed on the estimates are valid when the number of observations is large.

Keywords Nonlinear modeling · Variance modeling · Covariance structure · Predictions · Red spruce (*Picea rubens* Sarg.) · Balsam fir [*Abies balsamea* (L.) Mill.]

Communicated by Michael Köhl.

M. Fortin (✉) · J. Bégin
Département des sciences du bois et de la forêt,
Quebec, QC, Canada G1K 7P4
e-mail: mathieu.fortin@mrnf.gouv.qc.ca

J. Bégin
e-mail: jean.begin@sbf.ulaval.ca

G. Daigle
Département de mathématiques et statistique,
Université Laval, Quebec, QC, Canada G1K 7P4
e-mail: gaetan.daigle@mat.ulaval.ca

C.-H. Ung · L. Archambault
Natural Resources Canada, Canadian Forest Service,
Laurentian Forestry Centre, 1055 du P.E.P.S.,
P.O. Box 3800, Quebec, QC, Canada G1V 4C7
e-mail: Chhun-Huor.Ung@nrcan.gc.ca

L. Archambault
e-mail: Louis.Archambault@nrcan.gc.ca

M. Fortin
Direction de la recherche forestière,
Ministère des Ressources naturelles et de la Faune du Québec,
2700 Einstein, Quebec, QC, Canada G1P 3W8

Introduction

Historically, growth and yield models have been used either to synchronize inventory data or to provide growth forecasts. When using a stand-level approach, growth and yield models parameterized with permanent sample plot data aim at predicting future stand characteristics, such as basal area, stocking or stand volume (cf. Vanclay 1994, Sect. 2). However, from a statistical standpoint, repeated measurements taken on the same subject, for instance a plot, invalidate the basic assumption of independent error terms, which is at the basis of the traditional least squares methods (Steel et al. 1997, Sect. 7.10). Moreover, the

variances of the error terms often tend to increase along with increasing predicted values. The resulting heteroscedasticity represents a violation of another basic assumption of these traditional methods (Steel et al. 1997, Sect. 7.10). In cases of departures from the assumptions of independent error terms with homogeneous variances, the ordinary least squares (OLS) and the nonlinear least squares (NLS) methods still provide an unbiased estimator for the vector of parameters (Sullivan and Clutter 1972; Gregoire et al. 1995). However, the estimated variances of the parameter estimates are both biased and inconsistent (LeMay 1990).

Although some statistical tools are now available, many growth and yield models are still parameterized without considering these two statistical issues (e.g. Dhôte and Hervé 2000; Pretzsch et al. 2002; Fortin et al. 2003; Deleuze et al. 2004). As outlined in Gregoire et al. (1995), this is probably due to the complexity of error modeling and the traditional assumption that OLS and NLS estimators are unbiased. However, the biased variances resulting from such regressions hinder the selection of the appropriate explanatory variables (Gregoire et al. 1995). Somehow, the parameterization of a growth model with permanent sample plot data should take into account repeated measurements and heteroscedasticity in order to obtain unbiased statistical inferences.

Many authors have already addressed the matter of repeated measurements in growth and yield modeling by using either random effects and/or direct error modeling (e.g. Sullivan and Clutter 1972; Gregoire 1987; Goelz and Burk 1992; Gregoire et al. 1995; Beaumont et al. 1999; Fang and Bailey 2001; Hall and Bailey 2001; Nothdurft et al. 2006). Direct error modeling means estimating the covariances among the error terms, whereas random-effects specification assumes that the total model variance can be splitted into several variance components. Basically, both methods consist in parameterizing the variance-covariance (VC) matrix of the error terms. To limit the number of covariance parameters to be estimated, the design of the variance-covariance structure is typically assumed to be the same across the individuals (Vonesh and Carter 1992). Many VC structures have already been documented, e.g. compound symmetric, first-order autoregressive, and the moving average structures in direct error modeling (cf. Littell et al. 1996, p. 93; Pinheiro and Bates 2000, Sect. 5.3.3), or diagonally blocked, multiple of an identity, and general positive-definite in random-effects specification (Pinheiro and Bates 2000, Sect. 4.2.2). With permanent sample plot data, the main concern is often to choose a covariance structure that fits unequally spaced and unbalanced repeated measurements, which are common with this type of data. Gregoire et al. (1995) clearly demonstrated that it is possible to do so through a continuous-time autoregressive structure. However, this structure alone does not account for heterogeneous variances. Indeed, a transformation can be performed on the dependent variable to overcome the problem of heteroscedasticity (Steel et al. 1997, Sect. 9.16). Nevertheless, conversion problems arise when the resulting estimates and their inferences have to be retransformed to the original scale (Duan 1983). A weighted regression could also be performed as an alternative to transformations. However, choosing the appropriate weight requires some subjectivity and modeling the variance seems to be preferable, especially if the purpose of the model is to generate predictions and uncertainty assessments (Gregoire and Dyer 1989; Parresol 1993). In fact, unequally spaced repeated measurements and heteroscedasticity are common in forest growth modeling but rarely simultaneously considered in the common VC structures.

This study proposes a VC structure to take into account both problems in a context of growth modeling. This structure was tested on a nonlinear model. Basal area growth following a partial cutting was modeled with a von Bertalanffy–Richards equation (cf. Richards 1959). The model was parameterized with permanent sample plot data for two species in mixed stands: red spruce (*Picea rubens* Sarg.) and balsam fir [*Abies balsamea* (L.) Mill.]. The discussion focused on the effectiveness of the tested structure as well as its potential and limitations. The von Bertalanffy–Richards equation has been largely discussed in previous studies (e.g. Pienaar and Turnbull 1973).

Statistical developments

Statistical theory

According to the previously described context, a statistical nonlinear growth model parameterized with permanent sample plot data can be expressed as follows:

$$y_{ij} = f(\mathbf{x}_{ij};\boldsymbol{\beta}) + \varepsilon_{ij} \quad \text{for} \quad i = (1,2,\ldots,q), \quad j = (1,2,\ldots,m_i) \tag{1}$$

$$\boldsymbol{\varepsilon}_{i\bullet} = (\varepsilon_{i,1}, \varepsilon_{i,2}, \ldots, \varepsilon_{i,m_i})^{T} \quad \boldsymbol{\varepsilon}_{i\bullet} \sim N_{m_i}(\mathbf{0}, \mathbf{V}_i)$$

where y_{ij} is the dependent variable evaluated at measurement j of plot i, $\mathbf{x}_{ij}$ is vector of explanatory variables also evaluated at measurement j of plot i, $\boldsymbol{\beta}$ is a column vector of unknown parameters, $\boldsymbol{\varepsilon}_{i\bullet}$ is a column vector of m_i within-plot error terms associated with plot i, and T denotes a matrix transposition. The elements of $\boldsymbol{\varepsilon}_{i\bullet}$ are assumed to be normally distributed with mean $\mathbf{0}$ and variance-covariance $\mathbf{V}_i$.

Since many measurements were taken on the same plot, a serial correlation can reasonably be expected between the

errors associated with the i^{th} plot. Many structures already exist for modeling the within-subject covariance structure as previously mentioned. Among them, the time-continuous autoregressive structure is a generalization of the first-order autoregressive structure (Littell et al. 1996, p. 127; Pinheiro and Bates 2000, p. 236):

$$\mathrm{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) = \sigma^2 \rho^{|t_{ij} - t_{ij'}|} \quad (2)$$

where σ^2 is the residual variance, ρ is the correlation between two measurements spaced one unit of time apart, and $|t_{ij} - t_{ij'}|$ is the absolute distance in time units between measurements j and j' of plot i. The more distant from each other the measurements, the less correlated they are. The common first-order autoregressive structure is obtained by using the order of the measurements instead of time in Eq. (2).

The time-continuous autoregressive structure and its derivatives have been largely used in forest growth and yield modeling (e.g. Goelz and Burk 1992; Gregoire et al. 1995; Beaumont et al. 1999; Nigh et al. 2002). Being continuous in time, covariance structure (Eq. 2) no longer requires the measurements to be equally spaced and balanced. Its efficiency with permament sample plot data has already been demonstrated in Gregoire et al. (1995). However, due to the assumption of homogeneous variances through parameter σ^2, this structure does not account for heteroscedasticity under its current form. To take heteroscedasticity into account, the variance can be modeled through a function involving either the predicted values or some explanatory variables (Parresol 1993; Davidian and Giltinan 1995, Sect. 2.2.3; Littell et al. 1996, Sect. 8.2.4; Pinheiro and Bates 2000, Sect. 5.2). Variance modeling can be seen as a weighted regression whose weight is parameterized instead of being arbitrarily fixed. The power-of-the-mean function is a well-known example of a variance function based on predicted values:

$$\mathrm{Var}(\varepsilon_{ij}) = \sigma^2 \left| f(\mathbf{x}_{ij}; \boldsymbol{\beta}) \right|^{2\theta} \quad (3)$$

where σ^2 is the residual variance, $f(\mathbf{x}_{ij}; \boldsymbol{\beta})$ is the predicted value for measurement j of plot i, and θ is a parameter to be estimated. The power-of-the-mean variance function (Eq. 3) is a generalization of the constant coefficient of variation model, for which $\theta = 1$, and the Poisson-like variance structure, for which $\theta = 0.5$ (Davidian and Giltinan 1995, p. 23). To take into account unequally spaced repeated measurements and heteroscedasticity, the continuous-time autoregressive covariance structure (Eq. 2) and the variance function (Eq. 3) can be merged. The general form of the resulting function expresses the within-plot covariance:

$$\mathrm{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) = v_{ijj'} = \sigma^2 \left(f(\mathbf{x}_{ij}; \boldsymbol{\beta}) \right)^{\theta} \left(f(\mathbf{x}_{ij'}; \boldsymbol{\beta}) \right)^{\theta} \rho^{|t_{ij} - t_{ij'}|} \quad (4)$$

where $\sigma^2, \boldsymbol{\beta}, \theta$, and ρ are assumed to be constant across the plots.

When $j = j'$, Eq. (4) reduces to the original variance function (Eq. 3). From (Eq. 4), the VC matrix of within-plot error terms ($\mathbf{V}_i$) is obtained as follows:

$$\mathbf{V}_i = \begin{bmatrix} v_{i,1,1} & v_{i,1,2} & \cdots & v_{i,1,m_i} \\ v_{i,2,1} & v_{i,2,2} & \cdots & v_{i,2,m_i} \\ \cdots & \cdots & \cdots & \cdots \\ v_{i,m_i,1} & v_{i,m_1,2} & \cdots & v_{i,m_i,m_i} \end{bmatrix}. \quad (5)$$

Assuming a generalization of VC structure (Eq. 5) to all the plots, the overall VC matrix of the error terms ($\mathbf{V}$) encompasses q diagonal blocks, with each block $\mathbf{V}_i$ being the VC matrix of plot i. Usually, the plots are assumed to be independent from each other and the off-block elements within $\mathbf{V}$ are set to 0, such that

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_q \end{bmatrix}. \quad (6)$$

Estimation method and model diagnostics

Estimation methods for models with parameterized VC matrices are mostly based on a likelihood approach (Gallant 1987, Sect. 5.5; Davidian and Giltinan 1995, p. 36; Littell et al. 1996, p. 498; Pinheiro and Bates 2000, Sect. 2.2). To optimize the likelihood function with respect to the generalized VC structure, a generalized least squares (GLS) regression can be performed (Davidian and Giltinan 1995, Sect. 2.3.4). The generalized least squares algorithm consists in generating a preliminary OLS or NLS estimator for the vector of parameters. This is done by calibrating the model with no VC structure. Then, the resulting parameter estimates are kept constant in order to estimate the VC parameters. Model parameters are estimated again by keeping the VC parameter estimates constant, and so forth until the convergence of the likelihood function on a maximum. At convergence, a GLS estimator is obtained. Most subalgorithms for the optimization of model and VC parameters are tedious. We will not describe them here. Readers can refer to Gallant (1987, Sect. 1.4) for further details about the algorithms used for estimating model parameters. The maximization of the likelihood function with respect to VC parameters is usually performed with the expectation–maximization and/or the Newton–Raphson algorithms (cf. Lindstrom and Bates 1988; Wolfinger et al. 1994). These algorithms can be coded in any matrix

languages, such as IML in SAS system (SAS Institute Inc. 2002). The *gnls* function available in the S-PLUS or R software also performs such regressions (*cf.* Pinheiro and Bates 2000, Sect. 8.3.3; Pinheiro et al. 2004).

Once the VC parameters are estimated, the inversion of $\mathbf{V}$ provides the marginal contribution of each observation to the likelihood of the model. The inversed matrix $\mathbf{V}$, i.e. $\mathbf{V}^{-1}$, is required to obtain the maximum likelihood estimator of $\boldsymbol{\beta}$ and its VC matrix (Davidian and Giltinan 1995, Sect. 2.3.5):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}\mathbf{y} \tag{7a}$$

$$\mathrm{var}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \hat{\boldsymbol{\Omega}} = (\mathbf{X}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1} \tag{7b}$$

where $\mathbf{X}$ is the Jacobian matrix, i.e. a matrix of model partial derivatives with respect to $\boldsymbol{\beta}$, $\mathbf{y}$ is a column vector whose elements are y_{ij}, $\boldsymbol{\Omega}$ is the VC matrix of the parameter estimates, and the circumflex denotes the maximum likelihood estimator. A robust estimator of $\boldsymbol{\Omega}$ can also be obtained through sandwich estimators (Hall and Bailey 2001; McCulloch and Searle 2001, p. 212).

In fact, matrix $\mathbf{V}$ can be seen as a generalization for statistical models working with the assumptions of additive and normally distributed error terms. Under the assumptions of independent errors and homoscedasticity, matrix $\mathbf{V}$ is set to $\sigma^2\mathbf{I}$, with $\mathbf{I}$ being the identity matrix. A parameterization with $\mathbf{V} = \sigma^2\mathbf{I}$ reduces to a OLS or NLS regression according to the context.

Once the GLS estimator is obtained, the adequacy of the VC structure can be assessed with normalized residuals. A Cholesky decomposition of matrix $\mathbf{V}$ is performed to calculate the vector of normalized residuals ($\mathbf{r}_{\mathrm{norm}}$) as follows (Pinheiro and Bates 2000, p. 239):

$$\mathbf{r}_{\mathrm{norm}} = (\hat{\mathbf{C}}^{-1})^{\mathrm{T}}(\mathbf{y} - \hat{\mathbf{y}}) \tag{8}$$

where $\hat{\mathbf{y}}$ is a column vector, whose elements are the predicted values $f(\mathbf{x}_{ij}; \hat{\boldsymbol{\beta}})$, and $\hat{\mathbf{C}}$ is the upper triangle of the Cholesky decomposition of $\hat{\mathbf{V}}$. Normalized residuals can be seen as standardized residuals for models with parameterized VC matrix. Actually, the normalized residuals as shown in (Eq. 8) are not only weighted by their own variances, but also by their covariances. Consequently, if matrix $\mathbf{V}$ is properly parameterized, the normalized residuals should be independently and normally distributed with mean 0 and variance 1. Variance homogeneity is easily checked by plotting the normalized residuals against predicted values. Some statistical tests are also available (cf. Carroll and Rupert 1988, Sect. 3.4.3). The assumption of independent errors can be verified by calculating the empirical Pearson correlations among the normalized residuals. Then, the hypothesis that correlations are null is tested analytically with confidence intervals (cf. SAS Institute 2002). The semi-variogram technique also provides a measure of similarity among the within-plot normalized residuals (Gregoire et al. 1995).

In addition to diagnostic plots, likelihood ratio tests (LRT) can be performed to compare nested models that were fitted using the maximum likelihood method. A first model is considered nested in a second one if this second model is identical to the first except for the inclusion of a new explanatory variable or a new VC parameter. The statistic, which is computed as $2\log(L_2/L_1)$ with L being the likelihood of the model, follows a χ^2 distribution with $k_2 - k_1$ degrees of freedom, with k being the number of parameters. LRT are useful for assessing the significance of VC parameters (e.g. Fang and Bailey 2001; Hall and Bailey 2001). A significant probability associated with the statistic indicates that the additional parameter significantly improves the maximum likelihood of the model. Note that LRT based on restricted maximum likelihood values (REML) are valid only if the fixed-effects specification is the same for both models (Pinheiro and Bates 2000, p. 83).

Models parameterized with the same data can also be compared with the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) (Littell et al. 1996, Sect. 3.2.2; Pinheiro and Bates 2000, Sect. 2.4.1). Both statistics are based on log-likelihood values penalized for the number of parameters. AIC and BIC can be considered as ''parsimony'' criteria: the smaller the value of the statistics, the better the model in terms of fit and simplicity. Since the penalty for parameters is greater in the BIC, this criterion tends to be more conservative than the AIC.

Predictions

Once parameterized, the purpose of a model is usually to provide predictions. A general prediction theorem is presented in Valliant et al. (2000, Sect. 2.2). First, let us define the indices r and s as the unobserved and the observed part of their respective matrices or vectors:

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{ss} & \mathbf{V}_{sr} \\ \mathbf{V}_{rs} & \mathbf{V}_{rr} \end{bmatrix} \tag{9a}$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_s \\ \mathbf{X}_r \end{bmatrix} \tag{9b}$$

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_s \\ \mathbf{y}_r \end{bmatrix} \tag{9c}$$

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{\mathbf{y}}_s \\ \hat{\mathbf{y}}_r \end{bmatrix}. \tag{9d}$$

The elements of the observed part of the response vector, i.e. $\mathbf{y}_s$, are actually previous observations for some

particular plots, which can be used to obtain plot-specific responses. Henceforth, we will refer to these previous observations as ''prior information''. Let $\tilde{\mathbf{y}}_r$ be the empirical best predictor of the unknown part of the response vector, i.e. $\mathbf{y}_r$. This vector can be obtained as follows:

$$\tilde{\mathbf{y}}_r = \hat{\mathbf{y}}_r + \hat{\mathbf{V}}_{rs}\hat{\mathbf{V}}_{ss}^{-1}(\mathbf{y}_s - \hat{\mathbf{y}}_s) \quad (10)$$

where the last term represents the improvement when prior information is available. The prediction error associated with vector $\tilde{\mathbf{y}}_r$ has a variance that originates from two different sources of variability: the variability due to the error terms, which is defined in $\mathbf{V}$, and the variability due to the parameter estimates, which is encompassed in $\mathbf{\Omega}$. A third source of variability, which is the variability due to VC parameter estimates, has not been integrated into the calculation of prediction variances yet. The error variance of the empirical best predictor is

$$\begin{aligned}\mathrm{var}(\tilde{\mathbf{y}}_r - \mathbf{y}_r) = {} & \hat{\mathbf{V}}_{rr} - \hat{\mathbf{V}}_{rs}\hat{\mathbf{V}}_{ss}^{-1}\hat{\mathbf{V}}_{sr} \\ & + (\mathbf{X}_r - \hat{\mathbf{V}}_{rs}\hat{\mathbf{V}}_{ss}^{-1}\mathbf{X}_s)\hat{\mathbf{\Omega}}(\mathbf{X}_r - \hat{\mathbf{V}}_{rs}\hat{\mathbf{V}}_{ss}^{-1}\mathbf{X}_s)^{\mathrm{T}} \quad (11)\end{aligned}$$

where the last term represents the variance due to the parameter estimates, whereas the first two originate from the variance of the error terms. With the variance (Eq. 11), confidence intervals can be calculated around vector $\tilde{\mathbf{y}}_r$.

Equations (10) and (11) do not only apply to models with VC structures, but also to models with mixed effects. In fact, both approaches result in a parameterized matrix $\mathbf{V}$, which is then used to enhance the prediction. Note that the predicted values $\hat{\mathbf{y}}$ must be based on the fixed-effects parameters only when the theorem is used with a mixed model. In both cases, the best empirical predictor $\tilde{\mathbf{y}}_r$ and the variance $\mathrm{var}(\tilde{\mathbf{y}}_r - \mathbf{y}_r)$ respectively reduce to $\hat{\mathbf{y}}_r$ and $\mathbf{V}_{rr}$ when there is no prior information available for a particular plot.

The case study

The study site and the database

The 16-km^2 Lake Édouard Experimental Forest (LEEF) (46°45′N, 72°56′W) is located in the Laurentian section (L.4a) of the Great Lakes–St Lawrence River Region (cf. Rowe 1972). Created in 1918 for monitoring purposes, the LEEF has been part of La Mauricie National Park, Quebec, Canada, since 1970. Its climate is cool and humid. The nearby Shawinigan locality has a 1971–2000 mean annual temperature of 4.5°C (Environment Canada 2002). January is the coldest month with an average daily temperature of –13.2°C. The warmest month is July with an average of 19.6°C. The mean annual precipitation is 1,069 mm and the mean annual snowfall averages 249 cm. The LEEF average elevation is about 350 m above sea level. The topography is a succession of low-altitude hills interspaced with large areas of flat land (Heimburger 1941; Robitaille and Saucier 1998). The predominant soil type is a glacial till. The ecological features of the Laurentian section are similar to those of the well-known Acadian Forest Region (Rowe 1972). The spruce-fir-hardwood type described in Westveld (1953) is the most abundant forest type within the LEEF (Heimburger 1941; Ray 1956). It is usually located on lower slopes and well-drained flat lands (Heimburger 1941; Ray 1956; Grondin et al. 1996; Gosselin et al. 2001). Red spruce, balsam fir, yellow birch (*Betula alleghaniensis* Britton), and white birch (*Betula papyrifera* Marsh.) are the dominant species. Minor but common species are white spruce [*Picea glauca* (Moench) Voss.], eastern hemlock [*Tsuga canadensis* (L.) Carr.], eastern white pine (*Pinus strobus* L.), white cedar (*Thuja occidentalis* L.), red maple (*Acer rubrum* L.) and sugar maple (*Acer saccharum* Marsh.). Between 1950 and 1957, an experimental diameter-limit cutting was conducted at the LEEF. The treatment aimed (1) to avoid over-cutting the softwoods, (2) to favour spruce over fir, and (3) to reduce the cutting cycle from 40 to 20–30 years (Hatcher 1959). To meet the management goals, the diameter limits were set at 20 cm (8 in) for balsam fir and 41 cm (16 in) for red spruce and yellow birch at a stump height of 30 cm (12 in) (Ray 1956). Harvested volumes were estimated to 30–35% of the total volume of all stems larger than 8.9 cm (3.5 in) in diameter at breast height (dbh) (Ray 1956; Hatcher 1959).

The growth following this partial cutting was monitored with previously established permanent sample plots. These 405-m^2 (0.1-acre) plots had been systematically distributed all over the LEEF territory using a 201-m (660-ft) grid design. Within each plot, all the trees greater than 1.3 cm (0.5 in) in *dbh* had been tallied by 2.5-cm (1-in) diameter class. After the 1950–1957 partial cutting, successive measurements were done in 1956–1957, 1967, 1994–1996, and 2001–2004. Due to particular constraints, the last three measurements were partial, i.e. some plots were not visited or the data were unavailable.

A subsample of these permanent sample plots was used as a parameterization data set in order to model the basal area growth following the 1950–1957 cutting. The plots were selected among those for which we had at least one measurement since 1994 (203 plots). To limit the ecological variability as much as possible, the study focused on the spruce-fir-hardwood forest type for which the largest number of plots was available. The data set included 96 plots for a total of 239 measurements. Some characteristics

of the selected plots are summarized in Table 1 for the first measurement following cutting. The basal area development after logging is also shown for both species in Fig. 1.

The model

In southern Quebec, partial cutting is mostly regulated by merchantable basal area growth. Merchantable basal area is here defined as the basal area of all trees greater than 8.9 cm in diameter at breast height (dbh, 1.3 m). This variable was selected as the dependent variable for the modeling. Merchantable basal area was first plotted against time since logging to obtain a general trend of the data (Fig. 1), which supported the choice of a von Bertalanffy–Richards equation (cf. Richards 1959) as model. Expressed as a function of time since logging, the integral form of the equation is defined as follows:

$$y_{ij} = A_i\left(1 - \mathrm{e}^{B_i t_{ij}}\right)^{C_i} + \varepsilon_{ij} \quad (12)$$

where y_{ij} is the merchantable basal area at the measurement j of plot i ($m^2 \cdot ha^{-1}$), t_{ij} is the time since logging at measurement j of plot i (year), A_i, B_i, and C_i are the general parameters of the function, and ε_{ij} is the error term. Parameters may either be constant or vary according to some explanatory variables common to the measurements of plot i. Parameter A_i sets the asymptote of the equation, whereas parameters B_i and C_i change the rate of growth as well as the inflection point.

Table 1 Summary of the 1956–1957 measurement for the 96 permanent sample plots of the spruce-fir-hardwood forest type

Species and characteristics	Minimum	Mean	Maximum	Standard deviation
Red spruce				
Merchantable basal area[a] ($m^2 \cdot ha^{-1}$)	0.0	3.5	9.0	2.0
Merchantable density[a] (stems·ha^{-1})	0	236	1062	162
mdq250[b] (cm)	3.4	15.6	26.7	4.8
Density index ratio[c]	0.01	0.31	0.78	0.16
Balsam fir				
Merchantable basal area[a] ($m^2 \cdot ha^{-1}$)	0.2	2.2	7.1	1.6
Merchantable density[a] (stems·ha^{-1})	25	216	767	155
mdq250[b] (cm)	3.6	13.1	23.4	4.3
Density index ratio[c]	0.02	0.21	0.57	0.12
All species				
Merchantable basal area[a] ($m^2 \cdot ha^{-1}$)	3.9	12.4	31.2	4.9
Merchantable density[a] (stems·ha^{-1})	222	664	1408	276
Time since logging	1	3	6	1.9

[a] Merchantable: all trees greater than 8.9 cm (3.5 in) in diameter at breast height (dbh)

[b] mdq250: mean quadratic diameter of the 250 largest stems per hectare

[c] Density index ratio (DIR): as defined in Eq. (15)

Equation (12) must be modified to suit the partial cutting context. Actually, immediately after logging, i.e. at $t_{ij} = 0$, the residual merchantable basal area is likely to be different from 0. For the purpose of this study, an additional parameter was included in the model so that the intercept would be different from 0. Basically, two alternatives exist. This additional parameter makes it possible to move the curve either along the time axis (x-axis) or along the merchantable basal area axis (y-axis). In this case study, the first option was preferred, because it seemed more consistent from a biological point of view. The displacement along the time axis means that the residual merchantable basal area reduces to a displacement of general yield curves driven by some explanatory variables. In other words, this approach assumes (1) that species development on a particular site follows a family of yield curves and (2) that some post-harvest features related to the species actually indicate how developed the species is. The alternative, i.e. the translation along the y-axis, did not offer such biological consistency.

The translation along the x-axis was implemented by specifying an additional parameter (D_i) in the model:

$$y_{ij} = A_i\left(1 - \mathrm{e}^{B_i(t_{ij}+D_i)}\right)^{C_i} + \varepsilon_{ij}. \quad (13)$$

Previous works successfully related shifts of general yield curves to advance regeneration characteristics a few years after logging (Riopel 1999; Fortin et al. 2003).

After a few preliminary trials, the general parameters of Eq. (13) were set to

$$A_i = \mathrm{e}^{b_1 + b_2 \mathrm{DIR}_{i,1}} \quad (14a)$$

$$B_i = b_3 \quad (14b)$$

$$C_i = b_4 \quad (14c)$$

$$D_i = b_5 \mathrm{mqd250}_{i,1} - t_{i,1} + 1 \quad (14d)$$

where $\mathrm{DIR}_{i,1}$ is a density index ratio evaluated at first measurement of plot i, $\mathrm{mdq250}_{i,1}$ is the mean quadratic diameter of the 250 largest stems per hectare at the first measurement of plot i (cm), and $t_{i,1}$ is the time after logging for the first measurement of plot i. This latter variable acts upon the shift of the curve as a correction factor when the first measurement does not correspond to the first growing season following the diameter-limit cutting. This correction

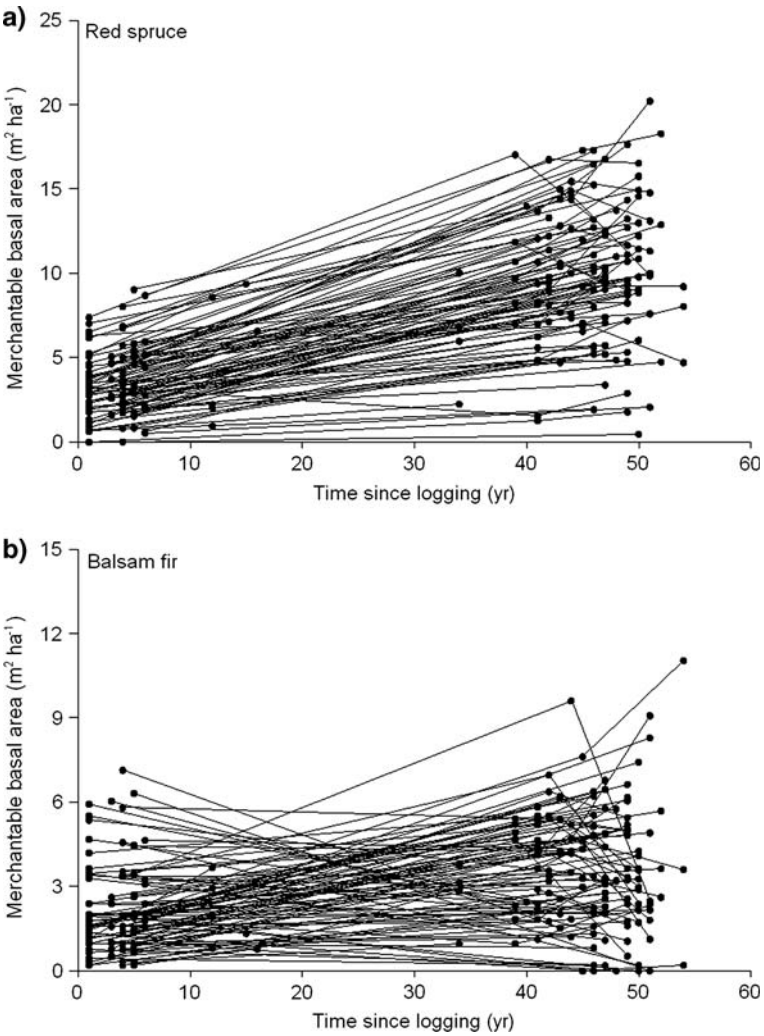

Fig. 1 Observed merchantable basal areas for both species (each *dot* represents a measurement, each *line* joins the measurements of a specific plot)

factor was necessary because the cutting was carried out from 1950 to 1957 depending on the plot, whereas the first measurement was done in 1956–1957. As a result, this first measurement does not match the immediate post-harvest conditions for 54 out of 96 plots.

The selection of these explanatory variables is based on information criteria (AIC and BIC). Other trials with additionnal covariates did not improve the fit of the model. As specified in Eq. (14a, d) the model assumes that the family of yield curves is driven by the density index ratio, whereas the shift along the x-axis is related to the mean quadratic diameter of the 250 largest stems.

The density index ratio (DIR) we used in this case study was computed as a ratio between the stand density index (SDI) calculated for the selected species and the all-species SDI. For convenience, Reineke's (1933) rule was used for calculating the different SDI, although there might be some variations at the species level (Pretzsch and Biber 2005). From this rule, the variable $\mathrm{DIR}_{i,1}$ is obtained as follows:

$$\mathrm{DIR}_{i,1} = \frac{\mathrm{DSP}_{i,1}}{\mathrm{DST}_{i,1}} \left(\frac{\mathrm{mqdsp}_{i,1}}{\mathrm{mqdst}_{i,1}} \right)^{1.605} \tag{15}$$

where DSP and DST, are respectively, the species and the all-species densities (stems·ha^{-1}), and mdqsp and mdqst are respectively the species and the all-species mean quadratic diameter (cm), with all these variables evaluated at the first measurements of plot i. This ratio is thought to be closely related to the site occupancy for a particular species.

Equation (13) with its parameters as defined in (Eq. 14a, b, c, d) was independently parameterized for red spruce and balsam fir. To take into account repeated measurements and heteroscedasticity, VC structure (Eq. 5) was included in the regression. Likelihood ratio tests (LRT) were performed to check the significance of VC parameters. Normalized residuals were also plotted to ensure the structure was adequate. Moreover, GLS predictions were compared with NLS predictions in order to illustrate the consistency and the reliability of the approach. Plot-specific responses

and their confidence intervals were computed with Eqs. (10) and (11), using prior information when available. For instance, predictions for the first measurement were computed with no prior information. The observed prediction error of the first measurement was then used to enhance the prediction for the second measurement, the errors on both the first and the second measurements were used to enhance the prediction for the third one, and so forth.

Results

Results of the likelihood ratio tests (LRT) and both AIC and BIC statistics are shown in Table 2. For both species, the AIC and the BIC clearly indicated that the best model was the one including VC structure (Eq. 5). LRT demonstrated the significance of the VC parameters θ and ρ. Adding variance function (Eq. 3) to a NLS-parameterized model significantly improved the likelihood of the model. Moreover, adding covariance structure (Eq. 2) over variance function (Eq. 3) also resulted in a significant improvement of the likelihood of the model.

For both species, the normalized residuals did not exhibit major departures from the assumption of normally distributed error terms with homogeneous variances. The observed correlations among the normalized residuals as well as the 1%-confidence intervals around the null hypothesis H_0 :Corr(ε_{ij}, $\varepsilon_{ij'}$) = 0 are illustrated in Fig. 2 for the three models of both species, i.e. the NLS model, the model with VF (Eq. 3) only, and the model with VC structure (Eq. 5). Some confidence intervals are quite large because there were very few observations for these particular lag distances. Normalized residuals of both the NLS model and the model with VF (Eq. 3) only exhibited a few significant correlations for some lags (Fig. 2a–d). VC structure (Eq. 5) seemed adequate since no observed correlations were significantly different from 0 at 1% after including the VC structure (Fig. 2e, f).

Parameter estimates of the ''best'' model, i.e. the model with VC structure (Eq. 5), are shown in Table 3. All model parameter estimates were significantly different from 0 at 5%, at least. Also, estimates of b_4 were significantly different from 1, meaning that the models could not be simplified by omitting this parameter. For both species, the estimates of b_2 indicated a positive effect of the density index ratio (DIR) on basal area growth. Increases in species density index with respect to the all-species density index resulted in a higher maximum basal area, i.e. a higher model asymptote. The estimates of b_5 also indicated a positive effect of mean quadratic diameter of the 250 largest stems per hectare (mdq250) on the shift of the yield curve along the time axis. As expected, the greater the mean quadratic diameter of the 250 largest stems per hectare, the greater the displacement of the curve.

The predictive abilities of each model are illustrated in Fig. 3 for both species. Plot 1 was selected as an example. Obviously, the NLS confidence intervals are the widest intervals, since the regression assumes homogeneous variances (Fig. 3a, b). On the other hand, models with variance function (Eq. 3) produce confidence intervals more consistent with the heteroscedastic pattern of the data (Fig. 3c, d). Note that OLS models and models with variance function (Eq. 3) lead to similar predicted values. The relevance of VC structure (Eq. 5) is clearly demonstrated in Fig. 3e and f. Using the available previous prediction errors makes it possible to obtain more accurate predictions and narrower confidence intervals, especially for the fourth measurement of each species. Note that VC structure (Eq. 5) can also be used without prior observations. In this case, it would generate mean predicted values that are similar to those of the two previous models [OLS and model with variance function (Eq. 3)].

Discussion

This case study demonstrates that VC structure (Eq. 5) may adequately account for heteroscedastic and correlated error

Table 2 Comparison of the goodness of fit for models with different variance-covariance (VC) features

Model	VC features	k[a]	AIC	BIC	Log-likelihood	LRT	Pr > χ^2
Red spruce							
1	None (NLS)	6	1031.04	1051.90	–509.52		
2	VF (Eq. 3) only	7	913.40	937.73	–449.70	1 vs 2	<0.0001
3	VC structure (Eq. 5)	8	882.92	910.73	–433.46	2 vs 3	<0.0001
Balsam fir							
1	None (NLS)	6	931.95	952.81	–459.97		
2	VF (Eq. 3) only	7	799.83	824.17	–392.92	1 vs. 2	<0.0001
3	VC structure (Eq.5)	8	790.15	817.96	–387.08	2 vs. 3	0.0006

LRT likelihood ratio test ; *VF* variance function

[a] k number of parameters

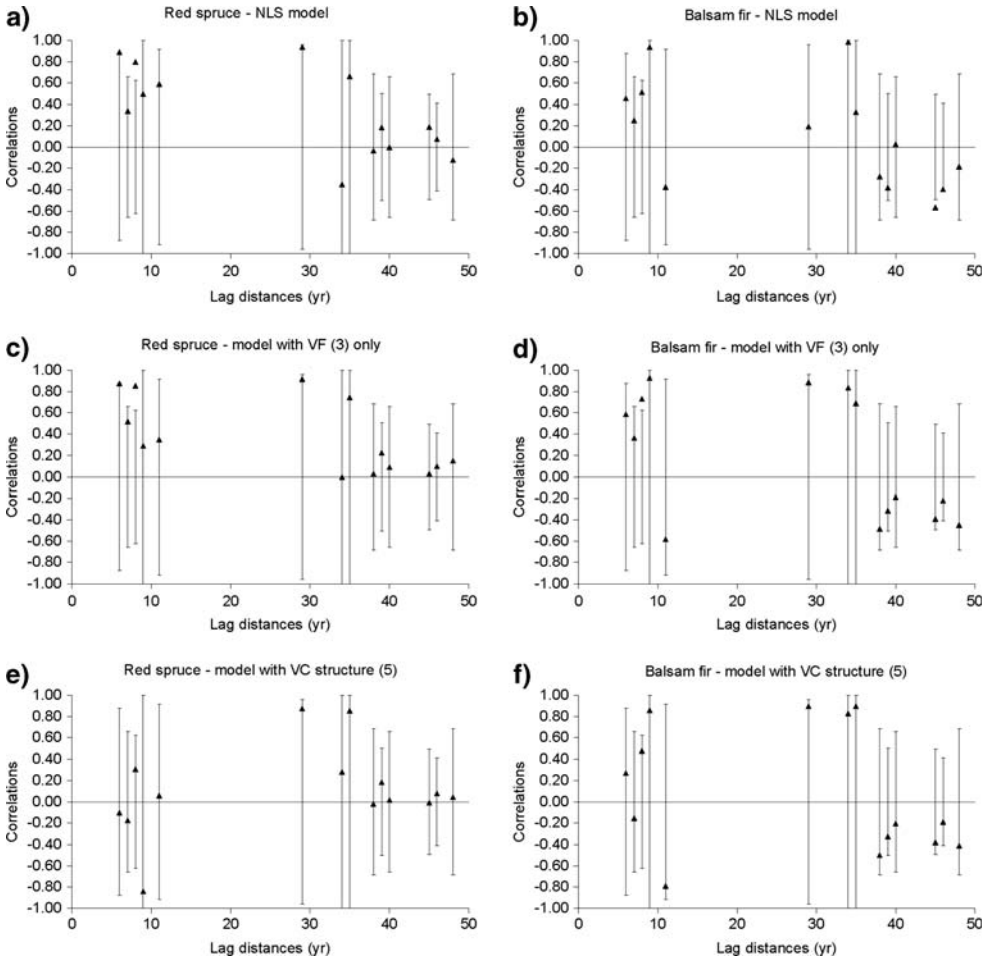

Fig. 2 Observed correlations among the within-plot normalized residuals and 1%-confidence intervals around the null hypothesis H_0 :Corr(ε_{ij}, $\varepsilon_{ij'}$) = 0 (the triangles indicate the observed correlations)

Table 3 Parameter estimates for model (Eq. 13) parameterized with VC structure (Eq. 5) (asymptotic standard errors are shown between parentheses)

Parameter estimates	Red spruce	Balsam fir
b_1	3.703 (0.556)**	1.092 (0.152)**
b_2 (Density index ratio effect)	0.4463 (0.1752)*	1.451 (0.364)**
b_3	-7.429×10^{-3} (3.249×10^{-3})*	-4.874×10^{-2} (1.874×10^{-2})**
b_4	2.376 (0.281)**	2.788 (0.382)**
b_5 (mdq250[a] effect)	3.348 (0.274)**	2.199 (0.660)**
θ (Variance parameter)	0.804	1.275
ρ (Correlation parameter)	0.944	0.900
σ^2 (Residual variance)	0.158	0.135

[a] *mdq*250: mean quadratic diameter of the 250 largest stems per hectare

*Significant at 5%;**significant at 1%

terms in a context of growth modeling using a stand-level approach. The goodness of fit of the model can be assessed through the normalized residuals, which are expected to be normally and independently distributed with homogeneous variances. For both species, there were no major departures from these assumptions and, consequently, the GLS estimator provided parameter estimates that were consistent and asymptotically normal with unbiased variances. If the number of observations is large, the approach makes it possible to select the explanatory variables according to their true significance level, a common problem with NLS regressions (Gregoire et al. 1995). Moreover, as the estimates and their standard errors are calculated from the original scale of the dependent variable, there is no need for a reverse transformation, which facilitates the interpretation of the inferences.

The parameterization of a VC matrix of the error terms (**V**) is a solution to departures from the assumptions of homogeneous variances and independent error terms. The approach proposed in this study is not the only way to

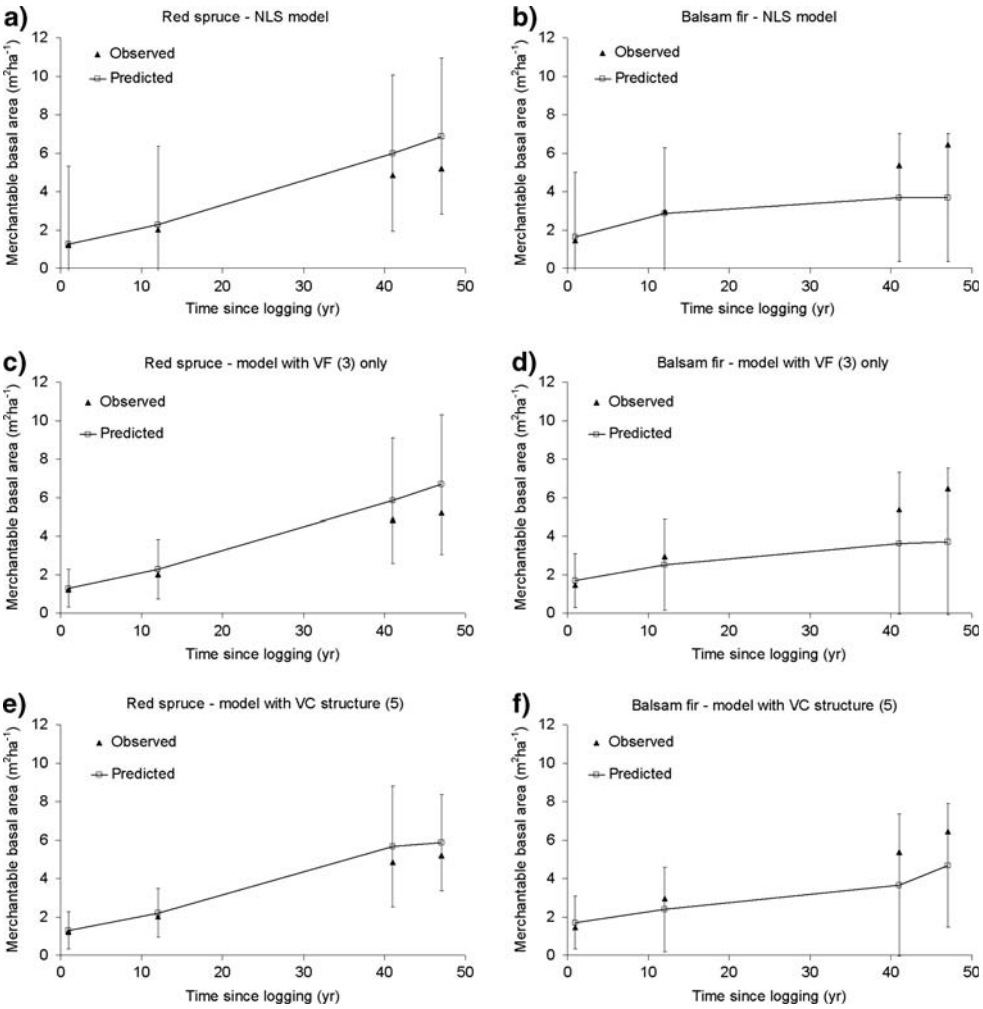

Fig. 3 Predictions and their 5%-confidence intervals for the different models of each species

parameterize this matrix. In fact, some simultaneously parameterized equations (cf. Gallant 1987, Sect. 6) and mixed-effects (both random and fixed) models (cf. Pinheiro and Bates 2000, Sect. 5.1.2) are additional statistical tools that make parameterization possible. In forestry, the hierarchical approach based on mixed models has been largely promoted in growth modeling (Gregoire et al. 1995; Fang and Bailey 2001; Fang et al. 2001; Hall and Bailey 2001). Basically, the hierarchical approach consists in specifying random effects associated with each level of grouping within the data set. The random effects of a given level are considered nested in the next level of grouping and the different levels are assumed to be independent. In the context of permanent sample plot data, inclusion of random effects in a growth model relies on the assumption that some unobserved influences act upon plot *i* in a constant fashion over time (Gregoire 1987). This constant influence is usually associated with fixed-effects parameters and is expressed as an error varying at the plot level. Many examples of multilevel mixed models can be found in Pinheiro and Bates (2000, Sect. 2.1.2).

In practice, a positive correlation can persist even after including random effects (e.g. Gregoire et al. 1995; Garber and Maguire 2003). In such cases, a continuous-time autoregressive covariance structure added to random effects gives better results (Gregoire et al. 1995; Fang and Bailey 2001; Garber and Maguire 2003). This study demonstrates that a VC structure alone may be sufficient to address the matters of heteroscedasticity and repeated measurements. In fact, the addition of random effects on the parameters of model (Eq. 13) was tested with and without VC structure (Eq. 5) and the resulting models exhibited lower likelihood values, nonsignificant LRT, as well as higher AIC and BIC statistics. These results were not surprising since there was no evidence of constant trends due to some unobserved plot factors among the empirical correlations shown in Fig. 2. In this case study, models with VC structure (Eq. 5) only were better models than those with random effects.

Actually, the two ways of parameterizing the VC matrix, i.e. VC structures and random effects, are not so different from each other. Even if random effects are referred to as

between-subject errors, their inclusion in a mixed model indirectly leads to the specification of a VC structure for the within-subject error terms (Littell et al. 1996, Sect. 3.2.2; Pinheiro and Bates 2000, Sect. 5.1.2). For linear models, Littell et al. (1996, p. 93) demonstrated that a random effect on the intercept and a compound symmetry VC structure yield exactly the same parameterization, with the same results and inferences. In practice, random effects and VC structures may interfere with each other in the model specification and the resulting VC matrix may be overparameterized (Pinheiro and Bates 2000, p. 204). Even if the likelihood converges on a maximum value, the model might not be the best one. For the lowest level of grouping, it is strongly recommended that this more complex model be compared with simpler models parameterized either with a VC structure only or random effects only. The comparison might reveal that a VC structure only is better than random effects or random effects with a VC structure. As outlined by Littell et al. (1996, p. 93), the contribution of the subject may be negligible in some repeated measures experiments.

Even if a model exhibits the lowest AIC and BIC statistics, there is no certainty that its VC parameterization adequately accounts for heteroscedasticity and correlated error terms. The only way to know whether the correction is appropriate is to plot the observed variances and correlations among the normalized residuals. Indeed, the lower the AIC and BIC statistics, the more likely the assumptions of independent error terms and homogeneous variances are to be expected. However, the selection of the final model should not be based on these statistics only. Moreover, the choice of the VC parameterization should take into account other factors besides the information criteria, such as the interpretability of the VC parameters (Pinheiro and Bates 2000, p. 409).

Gregoire (1987) interpreted plot random effects as influences due to variables mostly of an ecological nature, such as drainage and soil fertility. In this case study, the lack of constant correlations among the normalized residuals can be explained by the low variability of the ecological conditions, since all the plots belong to the same ecological group. Actually, random effects due to ecological variables are likely to be embedded in the effects of mortality and recruitment, which are thought to be far more important at the stand level. Except for the density index ratio (DIR), which can partially account for these effects, there is no variable that directly controls mortality and recruitment at the stand level. The main effect of having no control on these components is the disruption of the correlation between two measurements as the distance between them increases. As time goes on, basal area underestimations are likely to be offset by recruitment whereas mortality may compensate for overestimations. The continuous autoregressive covariance structure in VC structure (Eq. 5) clearly fits the disruption of correlation along with time.

In addition to correction for heteroscedasticity and correlated error terms, a VC structure can be used to improve model predictions. With mixed models, a similar enhancement is obtained by calculating a best unbiased linear predictor (BLUP) of random effects (cf. McCulloch and Searle 2001, Sect. 9), which is basically derived from the general prediction theorem in Valliant et al. (2000). In forestry, some studies have already demonstrated the improvement of the predictions by reinserting a BLUP into a mixed model (e.g. Fang and Bailey 2001; Hall and Bailey 2001; Nothdurft et al. 2006). Compared with previous work based on the hierarchical approach, the enhanced predictions in this case study are not as accurate. In fact, the accuracy of the predictions depends on the degree and the structure of correlation between the observations and the purpose of the model. Nothdurft et al. (2006) obtained very accurate subject-specific responses using two previous measurements. However, the variances of the random effects they specified in their model indicate that the error terms are highly correlated at the tree level. In this case study, the more distant in time the measurements, the lower the correlation between the observations. In these conditions, the best enhancements should be expected for measurements that are very close in time as shown in Fig. 3e, f. If our data structure had exhibited a higher degree of correlation, the prediction enhancement would have been greater.

Moreover, growth models are generally used to predict future forest conditions. Therefore, the future basal area measurements are unknown and the enhancement can be done only with prior information, i.e. with the observations prior to the prediction. This approach is the one we used in this case study. Indeed, more accurate enhancements would have been obtained if we had assumed that future measurements were available. Most studies based on mixed models performed the enhancement using this latter assumption. The BLUP are calculated with all the available observations, even observations subsequent to the prediction.

The VC structure proposed in this study does not consider the multivariate aspect of the analysis. Balsam fir and red spruce models were independently parameterized whereas they are likely to interact with each other. In fact, we could expect the error terms to be negatively correlated between the species. For the sake of the example, we decided to focus on the VC structure only. The species interactions can be handled through random effects as shown in Hall and Clutter (2004).

Conclusions

Progress in understanding the relationships between growth and environmental factors is often hindered by the lack of reliable and consistent statistical models. Most of the existing growth models ignore the impact of heteroscedasticity and unequally spaced repeated measurements on the reliability of prediction errors. In this study, a VC structure based on a variance function and a covariance structure has been proposed to take into account both statistical issues.

Although the hierarchical approach is strongly recommended in forestry, this case study demonstrates that a VC structure alone may provide the flexibility for simultaneously considering heteroscedasticity and correlated error terms. As a consequence, we recommend checking the assumption of constant influences due to the subjects before using random effects as proposed by the hierarchical approach. This can be done by comparing the AIC and BIC statistics for a mixed model and a model parameterized with a VC structure only. Plotting the normalized residuals can also give some hints about the appropriate correlation structure.

In addition to the flexibility it provides, the parameterization of a VC matrix also enhances the predictions. In forestry, this enhancement has already been demonstrated with mixed models (e.g. Fang and Bailey 2001; Fang et al. 2001; Hall and Bailey 2001; Nothdurft et al. 2006). In this study, we demonstrated it is also possible to improve the predictions through a VC structure.

Acknowledgments Funding for this work was provided by *le Fonds Québécois de la Recherche sur la Nature et les Technologies* (FQRNT), the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Forest Service (Laurentian Forestry Centre), Parks Canada (Albert van Dijk), *le Ministère des Ressources naturelles, de la Faune et des Parcs du Québec* and Abitibi-Consolidated Inc. (Jean Girard). Special thanks are due to Michèle Bernier-Cardou (Canadian Forest Service) for her comments on a preliminary version of the method, to Pamela Cheers and Isabelle Lamarre (Canadian Forest Service) for editing the text, to Bernard Parresol (USDA), Thomas Seifert (Technische Universität München), and four anonymous reviewers for their comments on an earlier draft of this manuscript, and to the people involved in field measurements. This work would not have been possible without the exceptional dedication of two senior foresters, R.G. Ray and R.J. Hatcher, who developed and conducted the long-term monitoring in the Lake Édouard Experimental Forest.

References

Beaumont J-F, Ung CH, Bernier-Cardou M (1999) Relating site index to ecological factors in black spruce stands: tests of hypotheses. For Sci 45(4):484–491

Carroll RJ, Ruppert D (1988) Transformation and weighting in regression. Monographs on statistics and applied probability 30. Chapman & Hall, New York

Davidian M, Giltinan DM (1995) Nonlinear models for repeated measurement data. Monographs on statistics and applied probability 62. Chapman& Hall, New York

Deleuze C, Pain O, Dhote J-F, Herve J-C (2004) A flexible radial increment model for individual trees in pure even-aged stands. Ann For Sci 61:327–335

Dhôte J-F, Hervé J-C (2000) Changements de productivité dans quatre forêts de chênes sessiles depuis 1930: une approche au niveau du peuplement. Ann For Sci 57:651–680

Duan N (1983) Smearing estimate: a nonparametric retransformation method. J Am Stat Assoc 78:605–610

Environment Canada (2002) Canadian climate normals or averages 1971–2000 (on line), The Green Lane, Environment Canada. Available at http://climate.weatheroffice.ec.gc.ca/climate_normals/index_f.html (reviewed May 10, 2004)

Fang Z, Bailey RL (2001) Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. For Sci 47(3):287–300

Fang Z, Bailey RL, Shiver BD (2001) A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. For Sci 47(4):550–562

Fortin M, Bégin J, Bélanger L (2003) Les coupes partielles: une alternative à la coupe à blanc dans les peuplements mixtes de sapin baumier et d'épinette rouge en termes de rendement. For Chron 79(5):948–956

Gallant AR (1987) Nonlinear statistical models. Applied probability and statistics series. Wiley, New York

Garber SM, Maguire DA (2003) Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. For Ecol Manage 179:507–522

Goelz JCG, Burk TE (1992) Development of a well-behaved site index equation: jack pine in north central Ontario. Can J For Res 22:776–784

Gosselin J, Grondin P, Saucier J-P (2001) Rapport de classification écologique du sous-domaine bioclimatique de l'érablière à bouleau jaune de l'est. Gouvernement du Québec, Ministère des Ressources naturelles, Direction de la gestion des stocks forestiers

Gregoire TG (1987) Generalized error structure for forestry yield models. For Sci 33(2):423–444

Gregoire TG, Dyer ME (1989) Model fitting under patterned heterogeneity of variance. For Sci 35(1):105–125

Gregoire TG, Schabenberger O, Barrett JP (1995) Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Can J For Res 25:137–156

Grondin P, Ansseau C, Bélanger L, Bergeron J-F, Bergeron Y, Bouchard A, Brisson J, De Grandpré L, Gagnon G, Lavoie C, Lessard G, Payette S, Richard PJH, Saucier J-P, Sirois L, Vasseur L (1996) Écologie forestière. In: Bérard JA, Côté M (eds) Manuel de foresterie. Les Presses de l'Université Laval, Sainte-Foy, pp. 133–279

Hall DB, Bailey RL (2001) Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. For Sci 47(3):311–321

Hall DB, Clutter M (2004) Multivariate multilevel nonlinear mixed effects models for timber yield predictions. Biometrics 60:16–24

Hatcher RJ (1959) Partial cutting with diameter limit control in the Lake Edward Experimental Forest, Quebec 1950 to 1956. Government of Canada, Department of Northern Affairs and National Resources, Forestry Branch, Forest Research Division, Project Q-44

Heimburger CC (1941) Forest site classification and soil investigation in the Lake Edward Forest Experimental Area. Government of Canada, Department of Mines and Resources, Lands, Parks and Forest Branch, Dominion Forest Service, Silvicultural Research Note No. 66

LeMay VM (1990) MSLS: a linear least squares technique for fitting a simultaneous system of equations with a generalized error structure. Can J For Res 20:1830–1839

Lindstrom MJ, Bates DM (1988) Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. J Am Stat Assoc 83(404):1014–1022

Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996) SAS system for mixed models. SAS Institute Inc., Cary

McCulloch CE, Searle SR (2001) Generalized, linear, and mixed models. Wiley, New York

Nigh GD, Krestov PV, Klinka K (2002). Height growth of black spruce in British Columbia. For Chron 78(2):306–313

Nothdurft A, Kublin E, Lappi J (2006) A non-linear hierarchical mixed model to describe tree height growth. Eur J For Res 125:281–289

Parresol BR (1993) Modeling multiplicative error variance: an example predicting tree diameter from stump dimensions in baldcypress. For Sci 39(4):670–679

Pienaar LV, Turnbull KJ (1973) The Chapman–Richards generalization of von Bertalanffy's growth model for basal area growth and yield in even-aged stands. For Sci 19(1):2–22

Pinheiro JC, Bates DM (2000) Mixed-effects models in S and S-PLUS. Springer, Heidelberg

Pinheiro JC, Bates DM, DebRoy S, Sarkar D (2004) *nlme*: Linear and nonlinear mixed effects models. R package version 3.1–53. The R Project for statistical computing (on line). Available at http://www.r-project.org/

Pretzsch H, Biber P (2005) A re-evaluation of Reineke's rule and stand density index. For Sci 51(4):304–320

Pretzsch H, Biber P, Ďurský J (2002) The single tree-based stand simulator SILVA: construction, application and evaluation. For Ecol Manage 162:3–21

Ray RG (1956) Site-types, growth and yield at Lake Edward Experimental Area Quebec. Government of Canada, Department of Northern Affairs and National Resources, Forestry Branch, Forest Research Division, Technical Note No. 27

Reineke LH (1933) Perfecting a stand-density index for even-aged forests. J Agric Res 46:627–638

Richards FJ (1959). A flexible growth function for empirical use. J Exp Bot 10:290–300

Riopel M (1999) Potentiel de croissance de la haute régénération préétablie de sapin baumier: une analyse dendrométrique au niveau de la tige et du peuplement. M.Sc. Thesis. Université Laval, Québec, Canada

Robitaille A, Saucier J-P (1998) Paysages régionaux du Québec méridional. Gouvernement du Québec, Ministère des Ressources naturelles

Rowe JS (1972) Les régions forestières du Canada. Environnement Canada, Service canadien des forêts, Publication no. 1300F

SAS Institute (2002) SAS OnlineDoc 9 (on line). SAS Institute Inc. Cary. Available at http://v9doc.sas.com/sasdoc/

Steel RGD, Torrie JH, Dickey DA (1997) Principles and procedures of statistics: a biometrical approach. McGraw-Hill, New York

Sullivan AD, Clutter JL (1972) A simultaneous growth and yield model for loblolly pine. For Sci 18:76–86

Valliant RL, Dorfman AH, Royall RM (2000) Finite population sampling and inference: a prediction approach. Wiley, Toronto

Vanclay JK (1994) Modelling forest growth and yield, applications to mixed tropical forests. CAB International, Wallingford

Vonesh EF, Carter RL (1992) Mixed-effects nonlinear regression for unbalanced repeated measures. Biometrics 48:1–17

Westveld M (1953) Ecology and silviculture of the spruce-fir forests of eastern North America. J For 51:422–430

Wolfinger R, Tobias R, Sall J (1994) Computing Gaussian likelihoods and their derivatives for general linear mixed models. SIAM J Sci Comput 15(6):1294–1310