

# Predicting individual tree mortality in northern hardwood stands under uneven-aged management in southern Québec, Canada

Mathieu FORTIN\*, Steve BÉDARD, Josianne DEBLOIS, Sébastien MEUNIER

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

(Received 11 May 2007; accepted 9 October 2007)

**Abstract** – This study proposes a generalized linear mixed model to predict individual tree mortality in northern hardwood stands under uneven-aged management. The model is based on a complementary log-log (CLL) link function, and was calibrated using permanent-plot data. Tree vigor, stem product, diameter at breast height and stand basal area were tested as explanatory variables. A plot and an interval random effect were specified to account for spatial correlations. When compared with the traditional logit link function, the CLL facilitates the inclusion of the time factor. In this case study, there was an important variability of mortality predictions between the plots and the intervals for a given plot. The interval random effect is thought to be associated with catastrophic mortality. Since both tree vigor and stem product proved to be significant mortality predictors, we recommend that these variables be evaluated to increase the accuracy of mortality models.

**mortality / generalized linear mixed model / random effects / hardwoods / single-tree selection cutting**

**Résumé** – Prédiction de la mortalité des tiges individuelles dans les peuplements de feuillus nobles sous aménagement inéquienne dans le sud du Québec, Canada. Cette étude présente un modèle linéaire mixte généralisé pour la prédiction de la mortalité dans les peuplements de feuillus nobles sous aménagement. Le modèle utilise une fonction de lien log-log complémentaire (LLC) et a été étalonné à l'aide de données de placettes permanentes. La vigueur de l'arbre, le produit, le diamètre à hauteur de poitrine et la surface terrière ont été testés comme variables explicatives. Des effets aléatoires de placette et d'intervalle ont été spécifiés dans le modèle afin de tenir compte des corrélations spatiales. Comparée au traditionnel logit, la fonction de lien LLC facilite l'inclusion du facteur temps. Dans ce cas d'étude, il existait une importante variabilité des prévisions entre les placettes et les intervalles de temps d'une même placette. On présume que l'effet aléatoire d'intervalle représente la mortalité catastrophique. Puisque la vigueur et le produit des tiges se sont avérés être des variables significatives, il est recommandé de les évaluer afin d'améliorer la précision des modèles de mortalité.

**mortalité / modèle linéaire mixte généralisé / effets aléatoires / feuillus nobles / jardinage par pied d'arbre**

## 1. INTRODUCTION

Uneven-aged management in hardwood stands aims at obtaining a sustained yield while maintaining a multicohort age structure over the long term [41, 52]. To meet these objectives, silvicultural treatments, such as single-tree or group selection cuttings, are usually carried out [41]. For the past two decades these silvicultural treatments have become increasingly popular in the province of Québec, Canada. From 1994 to 2003, selection cutting was carried out on about 40 000 ha of public lands annually [42–44]. This large-scale management is the result of the current provincial forest policy, which promotes uneven-aged silviculture in northern hardwood stands [39].

Despite the current regulation, there are still concerns regarding the success of this policy. In effect, mortality rates following selection cutting on public lands are double those observed on experimental study sites [3, 4]. Obviously, mortality losses reduce stand growth. Consequently, the cutting cycle is likely to be longer than the 15- to 25-year period that was initially planned [4].

Mortality is a complex phenomenon. Severe inter-tree competition gradually decreases the growth of individual trees and eventually leads to the death of some individuals [25]. The mixture of species and the structural complexity in northern hardwood stands make mortality modeling even more complex, because shade tolerance and life cycles differ among species. Moreover, in addition to normal competition-induced mortality, catastrophic mortality, which is thought to be due to extreme weather conditions as well as pest and disease outbreaks, may also occur [55]. It is also suggested that logging damage may result in increased mortality [41], especially among the smallest trees.

To get a better understanding of the phenomenon or to predict mortality rates, many authors have modeled mortality on a statistical basis (e.g. [8, 25, 46, 47, 49, 58]). Since the response variable is binary, logistic regressions are usually a convenient way to model mortality. However, common logistic regressions assume independent observations [34], an assumption that is unlikely to hold, since measured trees are usually grouped into plots. The traditional estimators still provide unbiased parameter estimates under departures from the assumption of independent observations [18, 27]. However, their

\* Corresponding author: mathieu.fortin@mrf.gouv.qc.ca

variances are biased, making the selection of the appropriate covariates difficult [18]. Recently, more complex estimators designed for dependent observations have been implemented into common statistical software. These estimators are mainly based on random effects specification (cf. [35]).

Since mortality predictions are central to growth forecasts in uneven-aged silviculture, a mortality model for individual trees is required to predict stand growth following selection cutting. The purpose of this study is to calibrate such a model for northern hardwood stands in Québec, Canada. More specifically, it aims to provide a model that will serve as a mortality module for a single tree-based stand simulator. It was calibrated with data from permanent plots that were established following experimental trials of selection cutting. To ensure the accuracy of the model predictions, a cross validation was carried out. The discussion focused on methodological issues related to mortality modeling.

## 2. STUDY SITES

### 2.1. Ecological features

From 1983 to 1999, the *Direction de la recherche forestière* (Forest Research Directorate) of the *Ministère des Ressources naturelles et de la Faune du Québec* (Québec Ministry of Natural Resources and Wildlife) carried out 44 experimental trials of selection cutting in northern hardwood stands. These experiments are confined to 20 study sites, which are distributed throughout the temperate forest zone. This zone roughly consists of a belt that runs in an east-west direction across the province between 45° 30' N and 48° 00' N.

Except for the St. Lawrence Plains, the soil deposits in this zone are mostly glacial tills. The topography is usually a succession of rolling hills interspaced with more-or-less flat lands [48]. The mean annual temperature ranges from 2.5 to 7.0 °C. The annual precipitation follows a geographical gradient that peaks at 1 200 mm in the most eastern parts of the province and slowly decreases to 800 mm in western Québec [48].

Despite the ecological variability, Québec's temperate forest zone is mostly dominated by mixed stands of hardwood species (cf. [19]). Sugar maple (*Acer saccharum* Marsh.) is usually the most abundant species in these stands, which develop on imperfectly to well-drained sites located on mid- and upper slopes. Yellow birch (*Betula alleghaniensis* Britton.) and American beech (*Fagus grandifolia* Ehrh.) are common associated species. In addition to these, white birch (*Betula papyrifera* Marsh.), red maple (*Acer rubrum* L.), American basswood (*Tilia americana* L.), American elm (*Ulmus americana* L.), balsam fir (*Abies balsamea* (L.) Mill.), red spruce (*Picea rubens* Sarg.) and white ash (*Fraxinus americana* L.) are frequently observed as well. An exhaustive list of the tree species observed in this study is given in Annex.

### 2.2. Single-tree selection cutting and monitoring

Single-tree selection cuttings were conducted on an experimental basis at each site. The objectives of the treatment were (i) to reduce mortality losses, (ii) to improve stand quality, (iii) to maintain an uneven-aged structure in the residual stand, and (iv) to be economically feasible on a 15- to 25-year cutting cycle. Non-vigorous trees were marked in priority for harvest. As much as possible, the marking was uniformly distributed throughout the stand, with special attention given to maintain a reverse J-shaped diameter distribution that was based on a particular residual basal area, *q* factor, and maximum diameter (cf. [31]). About 30% of the merchantable basal area, i.e., the total basal area of all trees having a diameter at breast height (dbh, 1.3 m) equal to or greater than 9.1 cm, was harvested. A selection cutting was carried out in one or several stands at each site. A part of each treated stand was left unmanaged as a control and a 0.5-ha permanent plot was established in both the unmanaged and the managed parts of the stand. The location of these plots is shown in Figure 1. In the managed part of a few stands, one or two additional plots were established. The purpose of these additional plots was to test different residual basal areas following selection cutting.

Within the plots, all trees with a dbh equal to or greater than 9.1 cm were numbered. The species was recorded and the dbh was measured using a diameter tape. Tree vigor and stem product were assessed following a two-class classification in both cases (cf. [31]). The tree vigor classification is done to identify the trees that are thought to be at highest risk of dying before the next cutting cycle. Basically, it is based on crown and bole defects. If more than one third of the crown showed evidence of dieback or damage, or if some defects could be observed, such as fungus, canker, decay or large open wounds, the tree was assigned to the non-vigorous class. Otherwise, it was considered as a vigorous tree. Likewise, depending on the straightness of the bole and external defects, broadleaved trees were classified as pulpwood or having sawlog potential. An exhaustive list of criteria used in both classifications can be found in Martel et al. [33]. It is worth mentioning that stem product was not evaluated for softwood species. After plot establishment, measurements were conducted on a five-year basis on average. In addition to diameter, dead trees and recruits were noted at each visit. One to four measurements are available for each plot depending on the date of the establishment. Altogether, 90 plots were available for the modeling of mortality rates. Their basal areas and densities are summarized in Table I.

## 3. MATERIAL AND METHOD

### 3.1. Database

In the context of a single tree-based stand simulator, Monserud and Sterba [38] mentioned that the simulator architecture determines how mortality must be calculated. Because the growth simulator was designed to work on a five-year basis in this case study, successive measurements of each per-

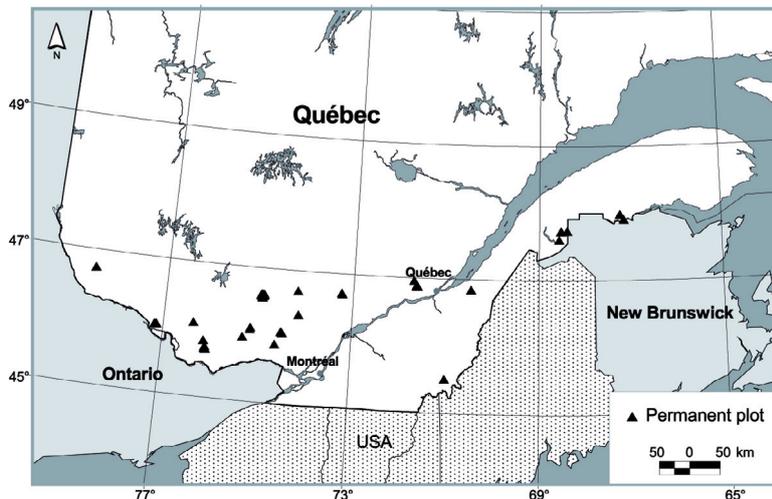


Figure 1. Location of the permanent plots that provided the calibration data.

Table I. Average basal areas and densities in harvested and control plots (standard deviation appears between parentheses).

Treatment and variable	Basal area (m <sup>2</sup> ·ha <sup>-1</sup> )	Density (st·ha <sup>-1</sup> )
After cutting		
Harvested (49 plots)	18.3 (2.2)	356 (71)
Control (41 plots)	25.8 (2.6)	467 (94)
5 years after cutting		
Harvested (49 plots)	19.6 (2.5)	368 (73)
Control (41 plots)	26.2 (2.9)	458 (92)
10 years after cutting		
Harvested (42 plots)	21.3 (2.9)	385 (71)
Control (34 plots)	27.1 (3.3)	458 (91)
15 years after cutting		
Harvested (24 plots)	22.7 (2.8)	402 (60)
Control (20 plots)	28.3 (3.0)	463 (99)

manent plot were paired in order to create nonoverlapping intervals. This way, the first measurement of each pair provided the initial conditions for a particular tree, whereas the second confirmed whether the tree had survived or died during the interval. Nonoverlapping intervals are also thought to reduce the degree of autocorrelation among the data [55].

The data set contained 22 different tree species. Some of these were scarce and so could not be individually taken into account in the model; consequently, they were grouped. The grouping was driven by factors such as species abundance and its ecology. The species groups are summarized in Table II. Altogether, the database contained 40 700 observations. A summary of the data is shown in Table III.

### 3.2. Statistical developments

For a single tree, death or survival can be represented as a binary variable that takes the value of 1 if the tree dies or 0 if

Table II. Species grouping for mortality modeling.

Group	Species
American beech	American beech
Yellow birch	Yellow birch
Red maple	Red maple
Sugar maple	Sugar maple
Other hardwood species	Black cherry, red oak, white ash, black ash, American elm, eastern hophornbeam, and American basswood
Other broadleaved species	White birch, trembling aspen, large-toothed aspen, pin cherry, striped maple, mountain ash
Balsam fir	Balsam fir
Other softwood species	White pine, eastern hemlock, red spruce, white spruce

it survives over a given time interval. Let  $y_{ijk}$  be this variable for tree  $j$  during interval  $k$  in plot  $i$ , such that  $i = (1, 2, \dots, p)$ ,  $j = (1, 2, \dots, q_i)$ , and  $k = (1, 2, \dots, r_i)$ . Note that the interval and the tree indices are both nested in the plots, while being crossed.

The aim of a mortality model is to predict the probability of a particular tree dying, which can be defined as  $Pr(y_{ijk} = 1) = \pi_{ijk}$ . Likewise, the probability of the same tree surviving is the balance of probabilities, i.e.,  $Pr(y_{ijk} = 0) = 1 - \pi_{ijk}$ . As mentioned in Rose et al. [49], the analysis of survival data in forestry is hindered by two major problems. First, the data are interval-censored because the exact time of death is unknown. (Actually, it is only known that a tree died during a particular interval, but not when it died during this interval.). Second, the assumption of independent observations is unlikely to hold, since the subjects, i.e., the trees, are nested in the plots.

**Table III.** Summary of the data set at the tree level (minimum diameter for all species was 9.1 cm).

Species group	$n^a$	dbh (cm)		Proportion (%)		
		Mean	Max	Vigorous <sup>b</sup>	Sawlog potential <sup>b</sup>	Dead <sup>c</sup>
American beech	5 143	22.9	71.2	66.3	35.8	4.7
Yellow birch	2 646	27.9	90.5	72.3	66.3	5.1
Red maple	974	24.8	63.1	54.9	39.9	9.4
Sugar maple	28 655	23.7	81.8	67.8	57.5	4.6
Other hardwood species	1 854	23.8	82.8	74.2	58.7	7.4
Other broadleaved species	234	16.6	45.0	60.7	22.6	36.3
Balsam fir	665	15.2	40.8	82.1		14.7
Other softwood species	529	24.0	64.8	84.5		7.2

<sup>a</sup>  $n$ : Number of observations, <sup>b</sup> at the beginning of the interval, <sup>c</sup> at the end of the interval.

Methods for considering interval-censored data are amply described in Cox and Oakes [9]. Let us define  $S(t_{ijk})$  as a survival function, where  $t_{ijk}$  is the time since the tree  $j$  in plot  $i$  entered the study. Thus, the survival probability  $1 - \pi_{ijk}$  over the interval  $(t_{i,j,k-1}, t_{ijk})$  is a marginal probability derived from  $S(t_{ijk})$ :

$$1 - \pi_{ijk} = \frac{S(t_{ijk})}{S(t_{i,j,k-1})} \tag{1}$$

Many survival functions  $S(t_{ijk})$  exist and we will not describe them here. Readers can refer to Cox and Oakes [9] for further details. For this study, we chose to express the survival function as an exponential function:

$$S(t_{ijk}) = e^{-\theta t_{ijk}} \tag{2}$$

where  $\theta$  is a parameter to be estimated. From equation (2), the survival probability becomes:

$$1 - \pi_{ijk} = \frac{e^{-\theta t_{ijk}}}{e^{-\theta t_{i,j,k-1}}} = e^{-\theta(t_{ijk} - t_{i,j,k-1})} \tag{3}$$

and a complementary log-log transformation on Equation (3) yields:

$$\ln(-\ln(1 - \pi_{ijk})) = \ln(\theta) + \ln(t_{ijk} - t_{i,j,k-1}) \tag{4}$$

Using a matrix notation, equation (4) can be reexpressed as a generalized linear model (GLM) based on a complementary log-log (CLL) link function:

$$\ln(-\ln(1 - \boldsymbol{\pi})) = \mathbf{X}\boldsymbol{\beta} + \ln(\Delta\mathbf{t}) \equiv \boldsymbol{\eta} \tag{5}$$

where  $\boldsymbol{\pi}$  is a vector whose elements are the  $\pi_{ijk}$ ,  $\mathbf{X}$  is a design matrix of explanatory variables,  $\boldsymbol{\beta}$  is a vector of unknown fixed-effect parameters, and  $\Delta\mathbf{t}$  is a vector whose elements are the  $t_{ijk} - t_{i,j,k-1}$ .

In model (5),  $\ln(\Delta\mathbf{t})$  can be considered as an offset, i.e., an independent variable whose parameter would be arbitrarily set to 1. Since Equation (5) can be rewritten as  $1 - \boldsymbol{\pi} = (e^{-e^{\mathbf{X}\boldsymbol{\beta}}})^{\Delta\mathbf{t}}$ , this offset variable also assumes that time is acting as a composite interest rate upon the annual survival probability. If the offset variable is, however, included in the design matrix  $\mathbf{X}$ , it becomes an explanatory variable. Thus, the complementary

log-log link function makes it possible to test this assumption of composite interest rate. If the assumption holds, the parameter estimate associated with the covariate  $\ln(\Delta\mathbf{t})$  should not be significantly different from 1.

The estimator for the vector of parameters ( $\boldsymbol{\beta}$ ) is based either on the maximization of a likelihood or a quasi-likelihood function. Usually, the maximum quasi-likelihood (MQL) estimator is preferred because it has important advantages over the pure maximum likelihood (ML) estimator (cf. [35]). Inference using the MQL method proceeds much as ML for the vector of parameters ([35]). Consequently, the MQL estimator also relies on the assumptions of homogeneous variances and independent error terms. Considering the data structure, this assumption is unlikely to be true, because the trees within the same plot are not spatially independent from each other. Furthermore, the mortality probabilities for trees observed during the same interval for a particular plot are likely to be correlated as well.

Most of the time, these uncontrolled effects are referred to as random effects. For several decades, mixed-effects (both random and fixed) models have become increasingly popular in forestry (e.g. [12, 13, 15, 18, 20, 26]). The mixed-models theory is amply described by some authors (cf. [29, 35, 45]), and it mostly applied to linear and nonlinear models.

Wolfinger and O'Connell [56] extended the mixed-models theory to generalized linear models by using a pseudo-likelihood approach, which is also referred to as a penalized quasi-likelihood (PQL). The inclusion of random effects in model (5) yields:

$$\boldsymbol{\eta} = \ln(\Delta\mathbf{t}) + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} \tag{6}$$

$$\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$$

where  $\mathbf{Z}$  is a design matrix associated with the random effect and  $\mathbf{u}$  is a vector of unobserved random effects.

In most generalized linear mixed models, the random effects are assumed to be normally and independently distributed with mean  $\mathbf{0}$  and variance-covariance  $\mathbf{G}$  (e.g. [6, 28, 35, 51]). Note that the assumed distributions for both the response and random effects can be different. The assumption of a normal distribution for the random effects facilitates the optimization of the penalized quasi-likelihood function.

The mathematical developments behind the optimization of the PQL function are tedious and we will not describe them

extensively. For further information, the reader can refer to Wolfinger and O'Connell [56] or to the documentation of the GLIMMIX procedure [51]. Basically, once the random and fixed effects are specified, a vector of pseudo-data ( $\mathbf{d}$ ) is generated from a first-order Taylor approximation as follows:

$$\begin{aligned} \mathbf{d} &= \hat{\boldsymbol{\eta}} + \Delta^{-1}(\mathbf{y} - \hat{\boldsymbol{\pi}}) \\ \Delta &= \frac{\partial \hat{\boldsymbol{\pi}}}{\partial \hat{\boldsymbol{\eta}}} \end{aligned} \quad (7)$$

where  $\Delta$  is a diagonal matrix whose elements are the first derivative of  $\boldsymbol{\pi}$  with respect to  $\boldsymbol{\eta}$ ,  $\mathbf{y}$  is a vector whose elements are  $y_{ijk}$ ,  $-1$  denotes a matrix inversion and the circumflex denotes the estimated values. Then, a linear mixed pseudo-model is defined from the pseudo-data using the standard mixed-models theory. The optimization of the PQL function provides preliminary estimators for the parameters. With these estimators, pseudo-data are recomputed and the PQL function is optimized again to produce more accurate estimates, and so forth. The algorithm runs until convergence of the parameter estimates, which produces the final PQL estimators of  $\boldsymbol{\beta}$  and  $\mathbf{G}$ . The GLIMMIX procedure in SAS System [51] enables the calibration of such complex generalized linear mixed models.

### 3.3. Model calibration and evaluation

Although they are not necessarily direct contributing factors [55], many variables may explain the mortality of individual trees. This is frequently related to tree size, stand density, individual tree competition and tree vigor [21]. In this case study, the species group (8 levels), the vigor class (2 levels), the product class (2 levels), and the dbh (cm) were identified as potential explanatory variables at the tree level. Distance-independent competition indices, such as the summarized basal area for all trees with dbh greater than the subject, the relative position of the subject tree in the cumulative basal area distribution, and the ratio between dbh and plot mean quadratic diameter, were tested in the model. These competition indices have been widely used in growth and mortality studies (e.g. [21,38,57]). At the plot level, the stand basal area ( $\text{m}^2\text{ha}^{-1}$ ), the tree density ( $\text{stems}\cdot\text{ha}^{-1}$ ), the duration of the time interval ( $y$ ) expressed in a logarithm form ( $\ln(\Delta t)$ ), and the treatment (2 levels) were also selected.

As regards the random effects, four hierarchal levels can be identified in this case study: the intervals nested in the plots, which are nested in the stands, which are in turn nested in the study sites. In addition to these four levels, the tree grouping represents another random effect, which is crossed with the interval grouping. All these random effects could be specified either on the intercept or in interaction with the explanatory variables.

A preliminary model was built without random effects. We used Akaike and Bayesian information criteria (AIC and BIC, see [45]) as indicator for the selection of the explanatory variables, even though we knew that these values were probably slightly biased because of the spatial- and time-correlations within the data structure. Some random effects were then specified in the model. Because the PQL estimator relies on

pseudo-data, it hinders using AIC and BIC values for random effects specification. Consequently, we used only the significance of the random effect variances as indicators when convergence could be reached. The significance of the variance components is usually tested at higher levels, such as  $\alpha = 10\%$  or  $\alpha = 30\%$  (e.g. [37]). We kept the random effects in the model if their variances were significant at  $\alpha = 30\%$ .

To test the validity of this preliminary model, a cross-validation was carried out. We used a resampling approach as suggested in Vanclay [55]. The model was calibrated 90 times, each time omitting one single plot in the data set. Then, predicted mortality probabilities were generated for each individual in the omitted plot. Because no prior information was available for the omitted plot, population-averaged predicted probabilities were computed. These predictions were obtained by integrating the conditional predictions with respect to the distributions of the random effects [35]. This integral has no closed-form solution, and consequently was solved numerically through Monte Carlo simulation.

Once the probabilities were generated for all the trees in the data set, the goodness of fit was assessed with a Hosmer-Lemeshow test. The statistic proposed by Hosmer and Lemeshow [23] follows a chi-square distribution. The observations are first ranked according to their predicted probabilities and then split evenly into ten groups. Predicted probabilities are averaged within each group and an expected number of events and non-events is calculated. The differences between observed and expected events and non-events are summed across the groups to produce the  $\chi^2$  statistic with 8 degrees of freedom. Large values of  $\chi^2$  indicate a lack of fit. Once the cross-validation provided adequate results, a final model was recalibrated using all the available data.

## 4. RESULTS

Preliminary trials revealed that none of the competition indices appeared to have a significant effect on mortality probabilities. Consequently, these variables were dropped out of the model. Several regressions were then performed to test the significance of the random effects in the model. Convergence could not be reached with all the levels of random effects, i.e., the interval, the tree, the plot, the stand and the site random effects. Actually, the estimation algorithm converged only when plot and interval random effects were specified on the model intercept. In this case study, the estimated variances of the random effects were significantly different from 0 at  $\alpha = 5\%$ . Consequently, both random effects were kept in the model.

As regards the fixed effects, a  $t$ -test was performed on a preliminary estimate for the parameter associated with  $\ln(\Delta t)$ . The estimate was significantly different from 0, but not significantly different from 1 ( $Pr > |t| = 0.7012$ ). Consequently, the model was reduced by specifying  $\ln(\Delta t)$  as an offset variable. All the other explanatory variables were significant at  $\alpha = 5\%$ , except for the plot density, which was dropped out of the model. The results of the  $F$ -tests performed on the fixed effects of the final model are shown in Table IV.

**Table IV.** Results of the calibration for the final model (using all the available data).

Fixed effects	df (numerator)	df (denominator)	F Value	Pr > F
Tree level				
Vigor class	1	40 678	744.03	< 0.0001
Product class	1	40 678	167.15	< 0.0001
Species	6	40 678	22.57	< 0.0001
dbh	1	40 678	15.24	< 0.0001
dbh <sup>2</sup>	1	40 678	31.47	< 0.0001
Species × dbh	7	40 678	10.47	< 0.0001
Plot level				
Treatment	1	106.5	17.67	< 0.0001
BA	1	134.0	15.69	0.0001
BA <sup>2</sup>	1	121.1	12.06	0.0007

dbh: Diameter at breast height (1.3 m), BA: basal area.

At the tree level, the mean mortality probabilities were significantly different between vigorous and non-vigorous trees. Likewise, there was a significant difference between pulp trees and stems with sawlog potential. The dbh significantly affected the mean mortality probabilities in a quadratic fashion, and the effect was significantly different between species. Moreover, harvested and control plots had significantly different mean mortality probabilities. The basal area had a significant quadratic effect on these mean probabilities as well.

To illustrate these differences, the model was used to generate mortality predictions. The parameter estimates that served for the predictions are presented in Table V. For the sake of the example, we used a reference tree. The differences for a particular explanatory variable are shown by changing the values of this variable while keeping the other covariates constant (Fig. 2). Tree vigor greatly affects the mean probability of mortality. In this example, the mortality rate for non-vigorous trees is five times higher on average (Fig. 2a). Pulp trees have mean mortality probabilities three times higher than trees with sawlog potential (Fig. 2b). The mean probability of mortality increases along with dbh for American beech and Balsam fir (Figs. 2c and 2d), while other species groups exhibit a u-shaped pattern, with mortality probabilities being higher among the smallest and largest stems. Note the different scales for the y-axis between Figure 2c and Figure 2d. Overall, balsam fir, other softwood species and other broadleaved species have higher mortality probabilities (Fig. 2d). The stems in the harvested plots had slightly lower mortality probabilities when compared with the control. Moreover, the effect of basal area on the mean predicted mortality probabilities exhibits an asymptotic trend, with the probabilities decreasing along with basal area increases.

Note that these differences are dependent on the characteristics of the reference tree. In fact, the model is not strictly linear because of the link function. The retransformation of equation (5) involves two exponentials. As a result, it is not possible to bring out the effect of a particular variable without considering the other covariates, because they affect each other during the retransformation. Consequently, the differences illustrated in Figure 2 might change proportionately with other

reference characteristics. Some trials, however, showed that these changes are relatively small. The mean predicted probabilities of mortality illustrated in Figure 2 are shown to help size the magnitude of the differences.

The above results are said to be population averaged, i.e., they were computed by integrating the conditional probabilities with respect to the distributions of the random effects. To illustrate the variability at both the plot and the interval levels, 95% confidence limits were calculated with the estimated variances of the random effects (Tab. V). These intervals were obtained by setting the random effects to  $\pm 1.96$  times their standard deviations, i.e., the square root of the variance. They are illustrated in Figure 3 for a particular combination of covariates. Basically, the variability due to the interval was slightly higher than the one associated with the plot. In both cases, the variability was non-negligible, as the upper bounds of the intervals exhibited predicted mortality probabilities that were at least twice those associated with the lower bounds. Again, because the model is not strictly linear, these differences might change slightly in proportion with other tree characteristics as references.

Once the cross-validation was done, the Hosmer-Lemeshow tests was carried out on the data. There was good adequacy between the mean predicted probabilities and the mean observed proportions for the ten groups of the test, as shown in Figure 4. The test produced a  $\chi^2$  statistic of 12.0 and an associated probabilities ( $Pr > \chi^2$ ) of 0.1533, which indicated that there was no major departure between the model predictions and the observed mortality rates.

## 5. DISCUSSION

The results of this study indicate that the individual mortality probability can be predicted with relative accuracy for an upcoming 5-year period in managed and unmanaged northern hardwood stands using a generalized linear mixed model. The extent of the calibration data set and the nature of the explanatory variables are two features contributing to the originality of this work. To our knowledge, there are no other studies addressing the issue of mortality modeling in managed hardwood stands with both a mixed-model approach and a data set as large as the one presented here.

The PQL estimator used in this mixed-model approach is thought to produce a consistent estimator for the vectors of parameters, though this question remains an open debate [6, 7, 28]. This complex estimation technique enables the specification of random effects in the model. In this case study, for instance, we managed to calibrate the model with both a plot and an interval random effect.

Many authors make the distinction between regular competition-induced mortality and more irregular catastrophic mortality (e.g. [25, 55]). Because catastrophic mortality is not clearly related to stand or site features, it is often difficult to distinguish this type of mortality with traditional logistic models. A plot random effect can account for unobserved sustained plot effects, such as aspect, drainage, soil density and composition, history and genetic origin of the growing stock [17].

**Table V.** Parameter estimates for the final model.

Fixed effect	Class	Parameter estimate	Pr >  t
Intercept		-0.0898	0.9275
Vigor class	Non-vigorous	1.5701	< 0.0001
	Vigorous	0	-
Product class (for hardwood only)	Pulp	0.9003	< 0.0001
	Sawlog potential	0	-
Species	American beech	-2.4472	< 0.0001
	Yellow birch	-0.8844	0.0111
	Red maple	-0.9159	0.0156
	Sugar maple	-1.3573	< 0.0001
	Other hardwood species	-0.3775	0.2889
	Other broadleaved species	0.7576	0.0608
	Balsam fir	0	-
	Other softwood species	0.1642	0.7126
dbh		0.02077	0.1914
dbh <sup>2</sup>		0.00057	< 0.0001
Species × dbh	American beech	-0.01974	0.2233
	Yellow birch	-0.06193	0.0002
	Red maple	-0.05533	0.0019
	Sugar maple	-0.05895	0.0002
	Other hardwood species	-0.07460	< 0.0001
	Other broadleaved species	-0.05567	0.0113
	Balsam fir	0	-
	Other softwood species	-0.05646	0.0043
Treatment	Harvested	-0.5410	< 0.0001
	Control	0	-
BA		-0.3049	0.0001
BA <sup>2</sup>		0.00557	0.0007
Random effect		Variance estimate	Pr >  z
Plot random effect		0.0464	0.0321
Interval random effect (nested in the plot)		0.1077	0.0001

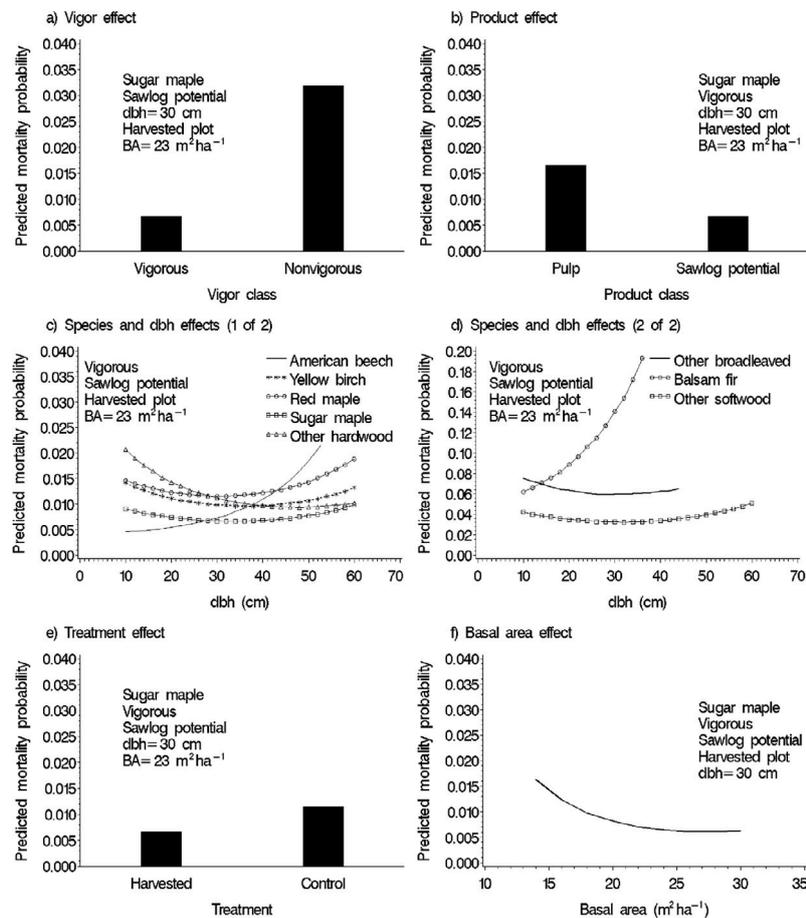
On the other hand, an interval random effect may encompass aggregate climatic influences [17]. For example, drought conditions, excessive rainfall, windstorms, pest and disease outbreaks, and decline for one or several years might increase mortality rates for a particular interval, independent of stand and site conditions. Therefore, an interval random effect could be seen as a crude estimate of catastrophic mortality, whereas the regular competition-induced mortality is partially taken into account through the plot random effect. In this case study, the interval random effect induced a slightly higher variability on mortality probabilities when compared with the plot random effect (Fig. 3). As the random effects are assumed to be independent from each other, significant increases of mortality probabilities are to be expected on poor site conditions submitted to episodic extreme climatic events.

The complementary log-log (CLL) link function we used also contributes to the originality of this study. Most studies on mortality use a logit link function (e.g. [1, 25, 46, 47, 53, 58]). Time is then inserted into the model to act as a power upon survival probabilities. Mathematically, the inclusion of the time factor makes sense. However, from a statistical standpoint, the logit link with the time factor makes the estimation process more complex, because it has to rely on a nonlinear adaptation of the generalized linear model (cf. [58]). A complementary

log-log link function has two advantages over the traditional logit link function. First, under the assumption of an exponential survival function, it facilitates the inclusion of the time factor, which can be done by specifying  $\ln(\Delta t)$  as an explanatory variable. Secondly, it makes it possible to test the assumption of time acting as a composite interest rate. If the assumption is valid, the parameter estimate associated with the explanatory variable  $\ln(\Delta t)$  should not be significantly different from 1. In other words, the CLL link function makes it possible to parameterize the survival probability (see Eq. (3)) as

$1 - \pi_{ijk} = \left( e^{-x_{ijk}\beta} \right)^{\Delta t_{ijk}^\delta}$ , where  $\delta$  is the parameter associated with  $\ln(\Delta t)$ . If the estimate for this parameter is not significantly different from 0 ( $\delta = 0$ ), there is no time effect. If the parameter estimate is not significantly different from 1 but significantly different from 0 ( $\delta = 1$ ), time acts as a composite interest rate over the survival probability. If the parameter estimate is significantly lower than 1 but higher than 0 ( $0 < \delta < 1$ ), the survival probability decreases slower than a composite interest rate. If the parameter estimate is significantly higher than 1 ( $\delta > 1$ ), the decrease is faster over time.

Basically, both link functions may yield a similar fit, but the CLL provides a more flexible and tractable model. In this case study, the effect of  $\ln(\Delta t)$  was significant, but it could



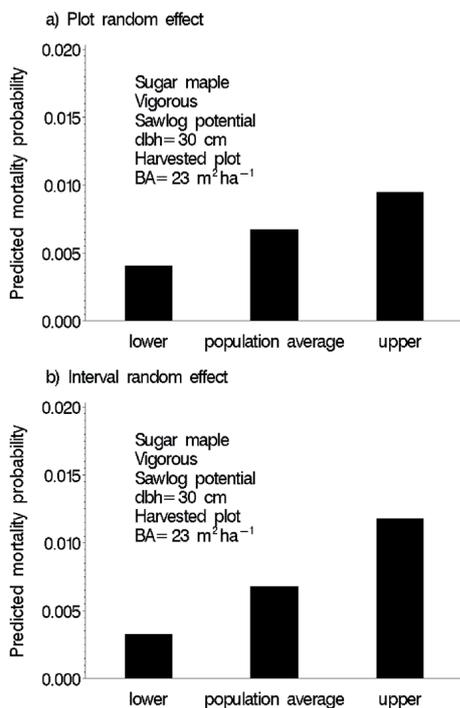
**Figure 2.** Effects of the explanatory variables on the mean predicted probabilities of mortality for an upcoming 5-year interval (other covariates are kept constant).

be specified as an offset because its parameter was not significantly different from 1. However, the value of  $\Delta t$  ranged from 4 to 7 years. Consequently, there is no guarantee that the assumption would hold for any longer time interval. The effect of  $\ln(\Delta t)$  could be both significantly different from 0 and from 1 for longer time intervals. In such a case, the assumption of time acting as a composite interest rate over survival probabilities would not hold anymore. This remains to be tested, and the complementary log-log link function we used in this paper might prove a convenient method for the test.

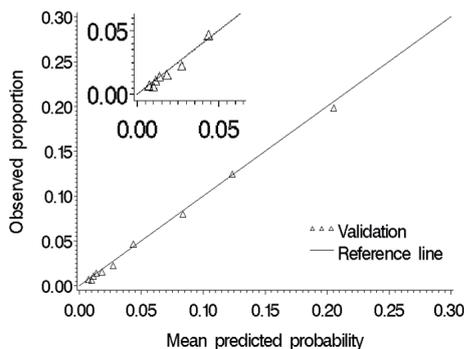
Although this study does not provide direct relationships among crown features, bole defects and mortality risk, it shows that using a two-class vigor classification based on these characteristics improves the accuracy of the predictions. This class variable exhibited the highest  $F$  value ( $F = 744.03$  in Tab. IV). In many cases, mortality models rely on past diameter increments for predicting mortality probabilities (e.g. [8,58]). Individual basal area growth has been found to be a good indicator of tree health [10]. In fact, the main advantage of using a vigor classification instead of past increments is that there is no need to wait until the third and fourth measurements to obtain the required explanatory variables, i.e., the past increments. On the other hand, this vigor classification remains

a crude indicator of the real tree vigor, since it relies on two classes only. Tree vigor is likely to be a gradient instead of a binary variable. Analyzing the incidence of particular defects and defoliation on mortality rates should lead to further improvements. Majcen et al. [32] already provide evidence that refinements of the vigor classification would enhance mortality predictions.

Hamilton [21] stated that mortality is frequently explained by (1) a measure of tree size, (2) stand density, (3) individual tree competition, and (4) growth rate (vigor). Variables from categories 1, 2, and 4 are already included in the model. However, the model does not include any direct measure of tree competition. The nonsignificance of the three competition indices that were tested in the model was unexpected. The plot size is probably an explanation of this result. The competition indices were all distance independent because tree coordinates were not available. Large plots, such as the ones in this case study, are likely to reduce the efficiency of these indices. A subject is actually compared with trees that may be located very far from the real competition area. The larger the plot, the more abundant are these “outsiders”. As a result, the local competition effect is drowned by the plot area. More complex distance-dependent indices might have yielded different



**Figure 3.** Example of a 95% confidence limits around the random effects for a reference tree.



**Figure 4.** Adequacy between the observed proportions of dead trees and the mean predicted probabilities of mortality for the ten groups of the Hosmer-Lemeshow test (the reference line indicates a perfect fit).

results, though they do not appear superior for predicting tree growth in these stands [30].

Another unexpected result is the significance of the stem product effect in the model ( $F = 167.15$  in Tab. IV). This effect indicates that trees with sawlog potential have better survival rates than pulp trees (Fig. 2b). Several tests were performed on the data set to check for possible correlations between the product class and the other explanatory variables. We found no evidence of such correlations, which could have biased the statistical inferences. The straightness of the bole and the absence of external evidence indicating stem decay are the two most important factors for classifying a particular tree in the sawlog potential class. Therefore, pulp trees might have a lesser me-

chanical resistance because of numerous limbs, a curved bole and internal stem decay. We speculate that this lesser resistance might induce a higher mortality rate for the pulp class.

Among the eight species groups, three had higher mortality probabilities: balsam fir, the other softwood group, and the other broadleaved group (Figs. 2c and 2d). Balsam fir is the species with the highest probabilities, which increases along with diameter. This pattern is in accordance with the species ecology. Balsam fir is usually short lived, though it is very shade tolerant [2, 14]. It is highly susceptible to windthrow, decay and spruce budworm [22, 36]. The mean predicted mortality probabilities indicate that large residual balsam fir stems are likely to die within a cutting cycle, no matter their vigor class (Fig. 2d).

The other softwood group and the other broadleaved groups encompass many species. As a result, it is difficult to point out a single contributing factor for each group. Most species of the other broadleaved group are known to be shade intolerant and short lived (e.g. [50]). The species of the other softwood group are subject to windfall [5, 40] or growing in suboptimal site conditions. These factors may explain the higher mortality rates for these species groups.

Although we do not consider other processes, such as recruitment and growth, the higher mortality probabilities for balsam fir, the other softwood and broadleaved species groups partially explain why these stands are dominated by hardwood species, especially sugar maple. On average, this latter species is the one with the lowest mean predicted probabilities of mortality. Yellow birch, red maple and other hardwood species have slightly higher mortality probabilities (Fig. 2c). Sugar maple is known to be very shade tolerant [16]. Actually, only the small merchantable stems of American beech exhibit comparable low mortality rates (Fig. 2c). This result is thought to be related to the high shade tolerance of American beech. However, among larger stems of American beech, the mortality probabilities are much higher (Fig. 2c). The province of Québec represents the northern boundary of the native range of this species. American beech is known to be sensitive to low winter temperatures, and can be severely damaged by late spring frosts [54]. These episodic low temperatures are common in the study areas, and large stems of American beech are likely to get damaged or even killed. Although beech bark disease is not currently considered as an important issue in Québec, some evidence indicates that the occurrence of the disease is increasing and it might have contributed to inflate the mortality rates.

Caspersen [8] already demonstrated the increase of annual mortality probabilities for small merchantable stems in the years following harvesting. In this case study, we have been unable to establish a direct relationship between time since cutting and the increase of mortality rates. In fact, the basal area effect is thought to encompass post-harvest mortality, as low basal areas are associated with the post-harvest stage in our data set. High mortality probabilities related to low basal area (Fig. 2f) might be partially due to post-harvest stress. As a result, it is impossible to know whether the source of this significant effect is basal area only or an interaction between basal area and treatment.

The same interaction also limits the interpretation of the effect of large basal areas over mortality rates. We expected mortality probabilities to increase along with increasing basal areas, because large basal areas may indicate higher competition-induced mortality. In this case study, however, mortality rates stabilize around 25–30 m<sup>2</sup> ha<sup>-1</sup>. These large basal areas are associated with control plots for which the treatment effect indicates higher mortality probabilities on average (Fig. 2e). Again, the treatment and the basal area effects cannot be accurately distinguished because they are closely associated in our data set. This represents a limitation to our study. The monitoring of unmanaged hardwood stands with low basal areas, as well as managed stands with large basal areas would have made it possible to distinguish the basal area effect from the treatment effect. Nevertheless, such stands are virtually impossible to find because the nature of the treatment necessarily implies basal area reductions. As a result, this problem is inherent to mortality modeling.

The treatment effect is unlikely to be the result of higher proportions of non-vigorous trees in the control plots, since the vigor was already accounted for in the model. In fact, there might be some spatial pattern in the occurrence of mortality. Windfall and limb breakage may damage, or even kill, neighboring trees. Since the control plots have higher densities and more non-vigorous trees, this collateral damage is likely to be more frequent when compared with harvested plots. The higher densities in the control plots may have increased the competition-induced mortality. This threshold of imminent competition mortality has been widely discussed as a conceptual part of management density diagrams [11, 24].

## 6. CONCLUSION

Recent statistical developments make it possible to consider the correlations among observations in the modeling of mortality. In this case study, we managed to calibrate a generalized linear mixed model for individual trees of eight species groups in both managed and unmanaged hardwood stands. Two random effects were specified in the model to account for the correlations. One of these random effects is thought to account for the aggregate climatic influence during the time interval, and could represent a crude assessment of the catastrophic mortality. Moreover, the complementary log-log link function that was specified in the model increases the mathematical tractability. It facilitates the integration of the time effect.

The statistical inferences drawn from the analysis indicate that both tree vigor and potential stem product have significant and important effects on mortality rates for individual trees in northern hardwood stands. The effects of both variables can be as important as those of standard covariates, such as tree diameter and stand basal area. This result outlines the requirements for vigor and product assessments in these stands in order to increase the accuracy of mortality predictions. Tree vigor can be seen as a substitute for past diameter increments in mortality modeling.

In common applications, mortality predictions are required for periods that exceed five years. As mentioned above, the mortality model in this study was designed to be part of a stand growth simulator, which is based on five-year steps. The simulator uses a Monte Carlo approach to simulate the evolution of tree features such as dbh, vigor and product classes over several steps, enabling mortality predictions for longer time intervals.

**Acknowledgements:** Special thanks are due to Daniel Mailly, Vincent Roy and François Guillemette (MRNFQ – Direction de la recherche forestière), and to two anonymous referees and the Associate Editor for their helpful comments on an early version of this paper. Jean Noël (MRNFQ – Direction de la recherche forestière) provided the illustration for Figure 1. Thanks to Gregory Crook for editing the text. The data used in this study would not have been available without the exceptional dedication of Zoran Majcen in establishing and conducting the selection cutting experiments throughout the province of Québec. Thanks to Laurier Groleau, Jocelyn Hamel, Pierrot Boulay, Étienne Boulay, and Jean-François Leblond (MRNFQ – Direction de la recherche forestière) for collecting and validating the data from these experiments.

## REFERENCES

- [1] Álvarez González J.G., Castedo Dorado F., Ruiz González A.D., López Sánchez C.A., von Gadow K., A two-step mortality model for even-aged stands of *Pinus radiata* D. Don in Galicia (Northwestern Spain), Ann. For. Sci. 61 (2004) 439–448.
- [2] Bakuzis E.V., Hansen H.L., Balsam fir, A monographic review, Univ. of Minnesota Press, Minneapolis, 1965.
- [3] Bédard S., Brassard F., Les effets réels des coupes de jardinage dans les forêts publiques du Québec en 1995 et 1996, Ministère des Ressources naturelles du Québec, Direction de la recherche forestière, Rapport, 2002.
- [4] Bédard S., Majcen Z., Meunier S., Coupe de jardinage dans les forêts feuillues du Québec – Mise à jour des résultats de recherche, InfoForêt 81 (2004) 3–6.
- [5] Blum B.M., Red spruce, in: Burns R.M., Honkala B.H. (Eds.), Silvics of North America, Vol. 1, USDA Agriculture Handbook No. 654, 1990, pp. 250–259.
- [6] Breslow N.E., Lin X., Bias correction in generalised linear mixed models with single component of dispersion, Biometrika 82 (1995) 81–91.
- [7] Browne W.J., Draper D., A comparison of Bayesian and likelihood-based methods for fitting multilevel models, Bayesian Analysis 1 (2006) 473–514.
- [8] Caspersen J.P., Elevated mortality of residual trees following single-tree felling in northern hardwood forests, Can. J. For. Res. 36 (2006) 1255–1265.
- [9] Cox D.R., Oakes D., Analysis of survival data, Chapman and Hall Ed., New York, 1984.
- [10] Duchesne L., Ouimet R., Morneau C., Assessment of sugar maple health based on basal area growth pattern, Can. J. For. Res. 33 (2003) 2074–2080.
- [11] Erdmann G.G., Oberg R.R., Fifteen-year results from six cutting methods in second-growth northern hardwoods, USDA For. Serv. Res. Pap. NC-100, 1973.
- [12] Fang Z., Bailey R.L., Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments, For. Sci. 47 (2001) 287–300.
- [13] Fang Z., Bailey R.L., Shiver B.D., A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects, For. Sci. 47 (2001) 550–562.

- [14] Frank R.M., Balsam fir, in: Burns R.M., Honkala B.H. (Eds.), *Silvics of North America*, Vol. 1, USDA Agriculture Handbook No. 654, 1990, pp. 26–35.
- [15] Garber S.M., Maguire D.A., Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures, *For. Ecol. Manage.* 179 (2003) 507–522.
- [16] Godman R.M., Yawney H.W., Tubbs C.H., Sugar maple, in: Burns R.M., Honkala, B.H., (Eds.), *Silvics of North America*, Vol. 2, USDA Agriculture Handbook No. 654, 1990, pp. 78–91.
- [17] Gregoire T.G., Generalized error structure for forestry yield models, *For. Sci.* 33 (1987) 423–444.
- [18] Gregoire T.G., Schabenberger O., Barrett J.P. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements, *Can. J. For. Res.* 25 (1995) 137–156.
- [19] Grondin P., Anseau C., Bélanger L., Bergeron J.-F., Bergeron Y., Bouchard A., Brisson J., De Grandpré, L., Gagnon G., Lavoie C., Lessard G., Payette S., Richard P.J.H., Saucier J.-P., Sirois L., Vasseur L., *Écologie forestière*, in: Bérard J.A., Côté M. (Eds.), *Manuel de foresterie*, Les Presses de l'Université Laval, Sainte-Foy, Québec, 1996, pp. 133–279.
- [20] Hall D.B., Bailey R.L., Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models, *For. Sci.* 47 (2001) 311–321.
- [21] Hamilton D.A. Jr., A logistic model of mortality in thinned and unthinned mixed conifer stands of northern Idaho, *For. Sci.* 32 (1986) 989–1000.
- [22] Hatcher R.J., *Mortality and regeneration following partial cutting of spruce-balsam fir-hardwood stands at Lake Edward, P.Q.*, Government of Canada, Department of Northern Affairs and National Resources, Forestry Branch, Forest Research Division, Project Q-44, 1959.
- [23] Hosmer D.W. Jr., Lemeshow S., *Applied logistic regression*, 2nd ed., John Wiley & Sons, New York, 2000.
- [24] Jack S.B., Long J.N., Linkage between silviculture and ecology: an analysis of density management diagrams, *For. Ecol. Manage.* 86 (1996) 205–220.
- [25] Jutras S., Hökkä H., Alenius V., Salminen H., Modeling mortality of individual trees in drained peatland sites in Finland, *Silva Fenn.* 37 (2003) 235–251.
- [26] Leites L.P., Robinson A.P., Improving taper equations of loblolly pine with crown dimensions in a mixed-effects modeling framework, *For. Sci.* 50 (2004) 204–212.
- [27] LeMay V.M., MSLS: a linear least squares technique for fitting a simultaneous system of equations with a generalized error structure, *Can. J. For. Res.* 20 (1990) 1830–1839.
- [28] Lin X., Breslow N.E., Bias correction in generalized linear mixed models with multiple components of dispersion, *J. Am. Stat. Assoc.* 91 (1996) 1007–1016.
- [29] Littell R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., *SAS System for Mixed Models*, SAS Institute Inc., Cary, NC, 1996.
- [30] Lorimer C.G., Tests of age-independent competition indices for individual trees in natural hardwood stands, *For. Ecol. Manage.* 6 (1983) 343–360.
- [31] Majcen Z., Richard, Y., Ménard, M., Grenier, Y., *Choix des tiges à marquer pour le jardinage d'érablières inéquiennes*, Guide technique, Ministère de l'Énergie et des Ressources du Québec, Direction de la recherche forestière, Mémoire No. 96, 1990.
- [32] Majcen Z., Bédard S., Meunier S., *Accroissement et mortalité quinze ans après la coupe de jardinage dans quatorze érablières du Québec méridional*, Ministère des Ressources naturelles et de la Faune du Québec, Direction de la recherche forestière, Mémoire de recherche, No. 148, 2005.
- [33] Martel J., Bergeron C., Demers G., Fortin Y., Hénaire F., *Méthode d'échantillonnage pour les suivis des interventions forestières*, Exercice 2001–2002, Ministère des Ressources naturelles du Québec, Direction de l'assistance technique, 2001.
- [34] McCullagh P., Nelder J.A., *Generalized linear models*, 2nd ed., Chapman & Hall/CRC, Monographs on Statistics and Applied Probability 37, New York, 1989.
- [35] McCulloch C.E., Searle S.R., *Generalized, linear, and mixed models*, John Wiley & Sons, New York, 2001.
- [36] McLintock T.F., Factors affecting wind damage in selectively cut stands of spruce and fir in Maine and northern New Hampshire, USDA For. Serv. Northeastern Forest Exp. Stn., Sta. Pap. No. 70, 1954.
- [37] Milliken G.A., Johnson D.E., *Analysis of messy data*, Vol. 1, *Designed experiments*, Van Nostrand Reinhold Company, New York, 1984.
- [38] Monserud R.A., Sterba H., Modeling individual tree mortality for Austrian forest species, *For. Ecol. Manage.* 113 (1999) 109–123.
- [39] MRNFP. *Manuel d'aménagement forestier*, 4e éd., Ministère des Ressources naturelles, de la Faune et des Parcs du Québec, Direction des programmes forestiers, 2003.
- [40] Nienstaedt H., Zasada J.C., White spruce, in: Burns R.M., Honkala B.H. (Eds.), *Silvics of North America*, Vol. 1, USDA Agriculture Handbook No. 654, 1990, pp. 204–226.
- [41] Nyland R.D., *Silviculture: concepts and applications*, 2nd ed., McGraw-Hill Ed, Toronto, 2002.
- [42] Parent B., Fortin C., *Ressources et industries forestières*, Portrait statistique édition 1999, Ministère des Ressources naturelles et de la Faune du Québec, Direction du développement de l'industrie des produits forestiers, 1999.
- [43] Parent B., Fortin C., *Ressources et industries forestières*, Portrait statistique édition 2001, Ministère des Ressources naturelles du Québec, Direction du développement de l'industrie des produits forestiers, 2001.
- [44] Parent B., Fortin C., *Ressources et industries forestières*, Portrait statistique édition 2005, Ministère des Ressources naturelles et de la Faune du Québec, Direction du développement de l'industrie des produits forestiers, 2005.
- [45] Pinheiro J.C., Bates, D.M., *Mixed-effects models in S and S-PLUS*, Springer/Verlag Ed., New York, 2000.
- [46] Pothier D., Mailly D., Stand-level prediction of balsam fir mortality in relation to spruce budworm defoliation, *Can. J. For. Res.* 36 (2006) 1631–1640.
- [47] González J.R., Trasobares A., Palahí M., Pukkala T., Predicting stand damage and tree survival in burned forests in Catalonia (North-East Spain), *Ann. For. Sci.* 64 (2007) 733–742.
- [48] Robitaille A., Saucier J.-P., *Paysages régionaux du Québec méridional*, Ministère des Ressources naturelles du Québec, 1998.
- [49] Rose C.E. Jr., Hall D.B., Shiver D.B., Clutter M.L., Border B., A multilevel approach to individual tree survival prediction, *For. Sci.* 52 (2006) 31–43.
- [50] Safford L.O., Bjorkbom J.C., Zasada J.C., Paper birch, in: Burns R.M., Honkala B.H. (Eds.), *Silvics of North America*, Vol. 2, USDA Agriculture Handbook No. 654, 1990, pp. 158–171.
- [51] SAS Institute, *The GLIMMIX Procedure*, Nov. 2005, (On line), SAS Institute, Cary, NC, 2005, Available at <http://support.sas.com/rnd/app/papers/glimmix.pdf> [reviewed May 24th, 2006].
- [52] Smith D.M., Larson B.C., Kelty M.J., Ashton P.M.S., *The practice of silviculture: Applied forest ecology* (9th Ed.), John Wiley & Sons, New York, 1997.
- [53] Teck R.M., Hilt D.E., Individual-tree probability of survival model for the Northeastern United States, USDA For. Serv. Res. Pap. NE-642, 1990.
- [54] Tubbs C.H., Houston D.R., American beech, in: Burns, R.M., Honkala, B.H., (Eds.), *Silvics of North America*, Vol. 2, USDA Agriculture Handbook No. 654, 1990, pp. 325–332.
- [55] Vanclay J.K., *Modelling forest growth and yield, applications to mixed tropical forests*, CAB International, Wallingford, UK, 1994.
- [56] Wolfinger R., O'Connell, M., Generalized linear models: a pseudo-likelihood approach, *J. Statist. Comput. Simul.* 48 (1993) 233–243.
- [57] Wykoff W.R., A basal area increment model for individual conifers in the northern Rocky Mountains, *For. Sci.* 36 (1990) 1077–1104.
- [58] Yao X., Titus, S.J., MacDonald, S.E., A generalized logistic model of individual tree mortality for aspen, white spruce, and lodgepole pine in Alberta mixedwood forests, *Can. J. For. Res.* 31 (2001) 283–291.

## ANNEX

**Tree species.**

The following 22 tree species were observed in the permanent plots of this study:

Latin name	Common name
<i>Abies balsamea</i> (L.) Mill.	Balsam fir
<i>Acer pensylvanicum</i> L.	Striped maple
<i>Acer rubrum</i> L.	Red maple
<i>Acer saccharum</i> Marsh.	Sugar maple
<i>Betula alleghaniensis</i> Britton.	Yellow birch
<i>Betula papyrifera</i> Marsh.	White birch
<i>Fagus grandifolia</i> Ehrh.	American beech
<i>Fraxinus americana</i> L.	White ash
<i>Fraxinus nigra</i> Marsh.	Black ash
<i>Ostrya virginiana</i> (Mill.) K. Koch	Eastern hophornbeam
<i>Picea rubens</i> Sarg.	Red spruce
<i>Picea glauca</i> (Moench) Voss	White spruce
<i>Pinus strobus</i> L.	White pine
<i>Populus grandidentata</i> Michx.	Large-toothed aspen
<i>Populus tremuloides</i> Michx.	Quaking aspen
<i>Prunus pensylvanica</i> L.f.	Pin cherry
<i>Prunus serotina</i> Ehrh.	Black cherry
<i>Quercus rubra</i> L.	Red oak
<i>Sorbus americana</i> Marsh.	Mountain ash
<i>Tilia americana</i> L.	American basswood
<i>Tsuga canadensis</i> (L.) Carr.	Eastern hemlock
<i>Ulmus americana</i> L.	American elm