

Evaluating a single tree-based growth model for even-aged stands against the maximum size-density relationship: Some insights from balsam fir stands in Quebec, Canada

Mathieu Fortin¹, Stéphane Tremblay² and Robert Schneider³

¹AgroParisTech, UMR 1092 LERFoB, 14 rue Girardet, 54042 Nancy, France, e-mail: mathieu.fortin.re@gmail.com

²Direction de la recherche forestière, Ministère des Forêts, de la Faune et des Parcs du Québec

³Faculté des sciences biologiques. Université du Québec à Rimouski

August 20, 2014

ABSTRACT

In this study, we addressed the issue of model evaluation when long-term monitoring data are unavailable or inappropriate. More specifically, we fitted a single tree-based growth model for pure even-aged balsam fir stands and we compared stochastic predictions with an existing maximum size-density relationship (MSDR). Growth trajectories for plots of different initial densities and diameter distributions were simulated over a 70-year period using 500 realizations for each combination of initial density-diameter distribution.

Plots with symmetric diameter distributions showed long-term predictions that were consistent with the existing MSDR, with only some realizations above the maximum density. In addition to competition-induced mortality, the model reproduced the senescence phase in which the trajectories diverge from the MSDR. For the even-aged stands that were simulated, this phase was initiated when the average tree volume reached $0.2\text{-}0.3\text{ m}^3\text{tree}^{-1}$, which roughly corresponded to a dbh (diameter at breast height, 1.3 m from the ground) between 19 and 23 cm. Initial asymmetric distributions resulted in trajectories that tended

to flatten out, showing some indications that the slope of the MSDR would be smaller for those stands.

Although it cannot be generalized, our case study shows that a simple single tree-based growth model with a distance-independent competition index and no stand density index can reproduce an existing MSDR. The match between long-term predictions and an existing MSDR strengthens the confidence in the biological behaviour of the model.

1 INTRODUCTION

In forestry, single tree-based models have become increasingly popular in the last two decades. Whereas the Prognosis (Stage 1973) and STEMS (Belcher et al. 1982) models were among the very first of their kind in the early 1980s, their number has dramatically increased so that it would be impossible to list them all today. Porté and Bartelink (2002) and Fontes et al. (2010) provided extensive lists of single tree-based models, to which the reader can refer.

Single tree-based models usually include at least a mortality submodel and a survivor increment submodel. These submodels are fitted on specific intervals of time (e.g., 5 or 10 years), and predictions for longer time intervals are obtained by reinserting the predictions into the model as many times as needed. A thorough evaluation of a single tree-based model is not straightforward since it includes many issues, both theoretical and empirical (Vanclay and Skovsgaard 1997). At the tree level, the different components of the model can be separately evaluated using classical statistical measures such as the average bias and the root mean square error. The model consistency can then be tested on single growth intervals, as was done by Pretzsch et al. (2002). However, this kind of evaluation does not ensure the consistency of plot-level predictions over longer time intervals.

Evaluating plot-level predictions in the long term can be particularly difficult for two reasons. First, long-term monitoring data are often nonexistent or insufficient. Secondly, even if such information is available, it is likely that either major natural or human-made disturbances will have occurred. These disturbances are difficult to reproduce in the simulations and they contribute a significant amount of uncertainty to the predictions. Using a classical evaluation based on bias and root mean square error can be misleading in such a context because the bias can originate from an improper simulation of the disturbances and not from the model itself.

When the plot-level evaluation cannot be supported by empirical data, it might be an alternative to test the biological behaviour of the model, as suggested by Oderwald and Hans (1993). The characteristic behaviour of even-aged stands is well

known as their production has been studied since the beginning of the 20th century and even before (e.g. Meyer 1929). Among other even-aged stand properties, the maximum size-density relationship (MSDR) has been addressed in many papers (e.g. Drew and Flewelling 1979; Jack and Long 1996; Monserud et al. 2004; VanderSchaaf and Burkhardt 2008). The MSDR relies on the assumption of the existence of a boundary beyond which any increase in average size must be compensated for by a decline in the number of individuals (Jack and Long 1996). On a log-log scale, the MSDR is assumed to be linear. Growth trajectories usually show an asymptotic trend with respect to this line with, eventually, a senescence phase that results in a departure from the relationship (cf. VanderSchaaf and Burkhardt 2008).

Balsam fir (*Abies balsamea* (L.) Mill) is one of the major commercial species in the northern United States and in Canada (Frank 1990). In the province of Quebec, Canada, regular balsam fir-dominated stands are common as a result of the last spruce budworm outbreak and clearcutting since the beginning of the 20th century (Grondin et al. 2013). In order to better understand growth dynamics in these stands following thinning, a long-term experiment was undertaken in 1968 (cf. Pothier 1998).

In this study, we used these long-term monitoring data to fit a single tree-based model designed for predicting the growth of pure and even-aged balsam fir stands. Based on a simple architecture, the model was designed to be used either in a stochastic or a deterministic way. However, the number of plots, the thinnings and a spruce budworm (*Choristoneura fumiferana* Clemens) outbreak prevented us from evaluating the long-term behaviour of the model against empirical data.

The biological behaviour of these stands had already been studied by Bégin et al. (2001) who fitted a MSDR for pure balsam fir stands using data from four other experiments. We took advantage of this previously published MSDR to check the biological consistency of long-term stochastic projections. In addition to this, we also tested the impact of the initial diameter distribution on growth trajectories. The discussion provides some insights into this evaluation compared to a previously published MSDR for balsam fir in Quebec, as well as the impact of the diameter distribution over the MSDR.

2 MATERIAL AND METHODS

2.1 STUDY AREA AND DATA

The data we used to fit the components of the single tree-based growth model were taken from an experiment carried out at Lac Matapedia, located in eastern Quebec, Canada (48°36'N, 67°31'W). This area is part of the balsam fir-yellow birch bioclimatic domain (Robitaille and Saucier 1998). It originates from a clearcut that was carried out either in the mid 1930s (Pothier 1998) or in the early 1960s (Thibault et al. 1995) according to the sources. Some stem analyses taken in 1994 showed no sharp increase in radial growth in the early 1960s, which seems to favour the mid 1930s as the official date of the harvest. The area is largely dominated by balsam fir. Red spruce (*Picea rubens* Sarg.), white spruce (*Picea glauca* (Moench) Voss), yellow birch (*Betula alleghaniensis* Britt.), white birch (*Betula papyrifera* Marsh.), pin cherry (*Prunus pensylvanica* L.f.), mountain ash (*Sorbus americana* Marsh.), eastern white cedar (*Thuja occidentalis* L.) and red maple (*Acer rubrum* L.) are the associated species in the study area.

The initial purpose of the experiment was to evaluate the effect of a late pre-commercial thinning on tree growth. To do this, 32 permanent plots were established throughout the study area in 1968. The plots measure 400 m² in area and all of the trees within a plot with a dbh greater than 1 cm were numbered. For each of the trees identified, the species was recorded and the dbh was measured to the nearest mm. Ten plots were left untouched as a control, whereas two thinning intensities were tested in the other 22 plots. The first measurement available was immediately after the pre-commercial thinning in 1968 when the stand was 33 years old (Pothier 1998). The plots were revisited in 1973, 1978, 1983, 1989, 1994, 2000 and 2006. If some tree heights had been measured in the inventories prior to 1994, they were unavailable. In 1994, some stem analyses were taken in order to determine the site index of some plots. From 2000, five study trees were randomly selected in each plot for height measurements.

Between 1973 and 1983, a spruce budworm outbreak of rare intensity occurred in the Northeast region (cf. Blais 1983). The damage was so extensive that it was decided to use pest control means in order to control the outbreak. Pesticides (aminocarb or fenitrothion) were sprayed in 23 of the 32 plots between 1978 and 1982, inclusively (Thibault et al. 1995; Pothier 1998). In the other nine plots, the damage was so intense that most trees were dead by 1989 and, consequently, these plots were abandoned afterwards. In 1994, a regular shelterwood cutting was carried out in nine of the 23 remaining plots. About one third of the stems were har-

vested which represented slightly more than a quarter of the basal area. The stems were felled and bunched manually. The evolution of the stand basal area and density is shown in Table 1.

Table 1. Stand density and basal area at the different measurement dates (minimum and maximum values appear in parentheses).

Measurement date	Number of plots	Density (trees ha ⁻¹)	Basal area (m ² ha ⁻¹)
1968	32	7495 (4025-15350)	37.1 (25.0-50.1)
1972	32	6278 (3875-11250)	42.1 (32.0-53.4)
1978	32	4945 (3075-8525)	44.3 (35.3-53.7)
1983	32	2930 (300-6550)	32.9 (4.9-51.0)
1989	32	1848 (0-4175)	27.9 (0.0-46.7)
1994			
Before shelterwood cutting	23	2317 (1375-3625)	41.1 (32.5-49.7)
After shelterwood cutting ^a	9	1614 (1075-2225)	30.2 (25.8-35.6)
2000	23	1983 (925-3200)	39.3 (25.5-52.5)
2006	23	1912 (950-4225)	41.5 (26.9-55.7)

^a Based solely on the nine plots that were harvested

For the modelling of mortality and survivor growth, the successive measurements were paired in order to create non-overlapping growth intervals of 5 or 6 years. The first measurement provides the initial conditions, whereas the second measurement shows whether or not the tree survived and provides its diameter increment if it did. Because some species were rather scarce, a grouping had to be carried out. Balsam fir accounted for 97% of the observations and it was the only species that was abundant enough to be considered by itself. We created three

species groups with the remaining species. Spruce species were grouped together, as were birch species. Finally, all the other marginal species formed the last group. Henceforth, we will refer to these species as balsam fir, *Picea* spp., *Betula* spp., and the other species. A summary of the dataset is provided in Table 2.

Table 2. Summary of the dataset (the minimum and maximum values appear in parentheses).

	Other species ^a	<i>Betula</i> spp.	<i>Picea</i> spp.	Balsam fir
Number of observations	158	413	414	32599
dbh (cm)	7.9 (1.2-22.5)	7.4 (2.2-24.1)	10.9 (1.5-32.1)	9.8 (1.0-31.5)
bal ^b (m ² ha ⁻¹)	34.8 (0.0-53.6)	34.7 (0.0-53.5)	23.2 (0.0-53.7)	28.1 (0.0-53.7)
Number of dead trees	28	53	70	7672
Average dbh increment (mm yr ⁻¹)	1.2 (-1.2-4.2)	0.8 (-1.2-6.3)	2.1 (-1.0-8.2)	1.2 (-1.4-8.8)

^a Other than *Betula* spp., *Picea* spp. and balsam fir

^b Basal area of all trees with dbh larger than the subject in the plot

2.2 GROWTH MODEL

The basic components of a growth models are mortality, survivor increment and recruitment (Vanclay 1994, p.8). There were very few recruits in the dataset, which was a direct consequence of the stand structure. Actually, only a few recruits were observed in the last two measurement campaigns of 2000 and 2006 as a result of the shelterwood cutting. This small number of recruits prevented us from fitting a recruitment submodel. We focused on the other two components, mortality and survivor increment, for which we had enough data. However, the absence of the recruitment component in the growth model limits its scope to even-aged stands.

The survivor increment was defined as the dbh increment during the growth intervals. We tested two sets of covariates: tree- and plot-level covariates. Tree species and dbh counted among the tree-level covariates, as well as a distance-independent competition index, the basal area of the trees with dbh larger than the subject in the plot. The plot-level covariates were the occurrence of spruce budworm

defoliation in the previous or the upcoming growth interval, the length of the interval (5 or 6 years), plot basal area and stem density, the type of pre-commercial thinning in 1968 and the occurrence of a shelterwood cutting in 1994.

The way the effects of the spruce budworm defoliation and pesticide sprays were considered in the submodels deserves further explanation. The 1973-1978 and 1978-1983 intervals were known to be affected by spruce budworm defoliation. However, for the 1978-1983 interval, Pothier's (1998) results showed that the effect of the spruce budworm outbreak was limited in the 23 plots that were protected between 1978 and 1982. For the sake of simplicity, we considered that the spruce budworm effect was negligible in those plots during the 1978-1983 growth interval.

Both submodels were evaluated using a 10-fold cross-validation. Prior to the statistical fitting, the plots were first divided into 10 groups and the submodels were then fitted 10 times, each time omitting one of the groups. For each fit, population-averaged predicted values were computed using the data from the plots that were not used for the fitting. After the 10 iterations, all the observations of the dataset had a predicted value, and some goodness-of-fit statistics were computed in order to evaluate each submodel.

2.2.1 Mortality submodel

A generalized linear model was used to model m_{ijk} , defined as a binary variable whose value is 1 in case tree j in plot i died during interval k , or 0 otherwise. The mortality submodel can be defined as follows:

$$Pr(m_{ijk} = 1) = 1 - e^{-e^{\mathbf{x}_{ijk}\boldsymbol{\beta} + b_{ik}}} \quad (1)$$

$$\mathbf{x}_{ijk}\boldsymbol{\beta} = \beta_0 + (\beta_{1,s} + \beta_{2,s}sbw_{ik-1})dbh_{ijk} + \beta_3dbh_{ijk}^2 + \beta_{4,s}bal_{ijk} + \beta_5\ln(\Delta t_{ik})$$

where sbw_{ik-1} is a dummy variable that takes the value of 1 if plot i suffered from spruce budworm defoliation in the previous growth interval $k - 1$, dbh_{ijk} is the diameter at breast height (cm) of tree j in plot i at the beginning of interval k , bal_{ijk} is the basal area (m^2ha^{-1}) of all trees with dbh larger than tree j in plot i at the beginning of interval k , Δt_{ik} is the duration of the time interval (yr), s is the species group index, such that $s = 1, 2, 3, 4$, and b_{ik} is a growth interval random effect, so that $b_{ik} \sim N(0, \sigma_{int}^2)$.

The above model is actually a generalized linear mixed-effects model based on the complementary log-log link function (cf. McCullagh and Nelder 1989). This link function offers greater mathematical tractability for the inclusion of time as

a compound interest effect (Fortin et al. 2008). This submodel accounts for the occurrence of a spruce budworm outbreak in the previous growth interval as well as asymmetric competition through the bal_{ijk} variable. Including the shelterwood treatment as well as the pre-commercial thinning treatments in the model caused some convergence problems. The effects of these variables were further investigated through Pearson residuals, which indicated that their magnitude was rather small.

Because the model is nonlinear and contains a growth interval random effect, the computation of population-averaged predictions is not straightforward. Actually, to obtain $E[1 - e^{-e^{x_{ijk}\beta + b_{ik}}}]$ for a particular tree, it is necessary to compute the integral of the predictions conditional on the random effect over the distribution of the random effect, i.e., $E[1 - e^{-e^{x_{ijk}\beta + b_{ik}}}] = \int (1 - e^{-e^{x_{ijk}\beta + b_{ik}}}) \varphi(b_{ik}; \sigma_{int}^2) db_{ik}$ where $\varphi(b_{ik}; \sigma_{int}^2)$ is the probability density function of a normal distribution with mean 0 and variance σ_{int}^2 . This integral has no closed-form solution and, consequently, some numerical integration method must be used. The Gauss-Hermite quadrature is one of the methods that can be used to obtain unbiased predictions, as reported in Fortin (2013). An example of this method is annexed to this paper (see Appendix A).

All the predictions provided by the cross-validation were corrected using the Gauss-Hermite quadrature in order to obtain population-averaged predictions. The area under the receiver operating characteristic curve (AUC), which is a well-known statistic for evaluating such generalized linear models (cf. Lasko et al. 2005), was calculated from these population-averaged predictions. The final version of this submodel was fitted with the whole dataset.

2.2.2 Diameter increment submodel

The diameter increment submodel is used to predict the average annual dbh increment ($\text{mm}\cdot\text{yr}^{-1}$) for the upcoming 5- or 6-year interval, defined as y_{ijk} . Visual checks on the distribution of y_{ijk} revealed that it was normally distributed. Consequently, we used a linear mixed-effect model that could be expressed as:

$$y_{ijk} = \mathbf{w}_{ijk}\boldsymbol{\gamma} + b_{ij}dbh_{ijk} + \varepsilon_{ijk} \quad (2)$$

$$\mathbf{w}_{ijk}\boldsymbol{\gamma} = \gamma_0 + \gamma_{1,s} + \gamma_2dbh_{ijk} + \gamma_3dbh_{ijk}^2 + \gamma_{4,s}sbw_{ik} + \gamma_5bal_{ijk}$$

where sbw_{ik} is a dummy variable that takes the value of 1 if plot i suffered from spruce budworm defoliation in the upcoming growth interval k , b_{ij} is a tree random effect, so that $b_{ij} \sim N(0, \sigma_{tree}^2)$, and ε_{ijk} is a residual error term, so that

$\varepsilon_{ijk} \sim N(0, \sigma_{res}^2)$. Note that the interaction between tree dbh and species was tested in the model but was non significant. Likewise, the effects of the pre-commercial thinning and the shelterwood cutting were found to be non significant once the bal_{ijk} variable was included and, consequently, these effects were not kept in the submodel.

2.2.3 Evaluation of short-term plot-level predictions

Before checking the behaviour of the model with respect to long-term projections, it should be tested whether the projections on a single growth interval are consistent. To do this, we selected the basal area (m^2ha^{-1}) and the stem density (tree- ha^{-1}) of plot i at the end of interval k as reference variables, defined as G_{ik} and N_{ik} , respectively. On the basis of submodels 1 and 2, the expectations of G_{ik} and N_{ik} can be derived as:

$$E[G_{ik}] = E \left[\sum_j e^{-e^{x_{ijk}\beta + b_{ik}}} \frac{25\pi(10 \cdot dbh_{ijk} + \Delta t_{ik}(\mathbf{w}_{ijk}\boldsymbol{\gamma} + b_{ij}dbh_{ijk} + \varepsilon_{ijk}))^2}{4000000} \right] \quad (3a)$$

$$E[N_{ik}] = E \left[\sum_j e^{-e^{x_{ijk}\beta + b_{ik}}} \right] \quad (3b)$$

where $e^{-e^{x_{ijk}\beta + b_{ik}}}$ is actually the balance of probability from the mortality sub-model, i.e., the probability of survival, and $10 \cdot dbh_{ijk} + \Delta t_{ik}(\mathbf{w}_{ijk}\boldsymbol{\gamma} + b_{ij}dbh_{ijk} + \varepsilon_{ijk})$ is the dbh of tree j in plot i at the end of growth interval k . The factor 25 at the numerator is an expansion factor to obtain per hectare values, whereas the factor 4000000 at the denominator is required for the conversion from mm^2 to m^2 .

Using 3a and 3b, it can easily be shown that (see Appendix A):

$$E[G_{ik}] = \sum_j E \left[e^{-e^{x_{ijk}\beta + b_{ik}}} \right] E[g_{ijk}]$$

$$E[N_{ik}] = \sum_j E \left[e^{-e^{x_{ijk}\beta + b_{ik}}} \right]$$

where

$$E[g_{ijk}] = \frac{25\pi}{4000000} \left(100 \cdot dbh_{ijk}^2 + 20 \cdot dbh_{ijk} \Delta t_{ik} \mathbf{w}_{ijk} \boldsymbol{\gamma} + \Delta t_{ik}^2 (\mathbf{w}_{ijk} \boldsymbol{\gamma})^2 + \Delta t_{ik}^2 (\sigma_{tree}^2 dbh_{ijk}^2 + \sigma_{res}^2) \right)$$

As mentioned in Section 2.2.1, the term $E \left[e^{-e^{x_{ijk}\beta + b_{ik}}} \right]$ is not straightforward here because of the interval random effect in $x_{ijk}\beta$. The integral of the conditional expectation over the distribution of the random effect needs to be calculated in order to obtain a population-average prediction of survival probability. This was done using the Gauss-Hermite quadrature numerical integration method (cf. Fortin 2013). The expectations of the basal area and stem density on single growth intervals were compared to the observations in terms of bias and relative bias.

2.3 MODEL IMPLEMENTATION AND LONG-TERM PREDICTIONS

In order to facilitate long-term growth simulations, the model was implemented in the CAPSIS platform (cf. Dufour-Kowalski et al. 2012). To obtain volume predictions, two other components were added to the model: a height-diameter relationship (Fortin et al. 2009b) and Honer et al.'s (1983) total volume equations. Note that the height-diameter relationship in Fortin et al. (2009b) is a general relationship derived from more than 100 000 height records. Considering the scarcity of height measurements in our data, it seemed preferable to use this relationship instead of fitting a new one. Also, it is worth mentioning that we used the total volume here and not the commercial volume because Bégin et al.'s (2001) MSDR is expressed in terms of mean total tree volume against stem density. The implementation of the different model components was done in such a way that the simulator could be used either in a stochastic or a deterministic mode. We took advantage of this feature to use the stochastic mode in our long-term simulations.

The stochastic simulations are based on the well-known Monte Carlo technique. All known sources of uncertainty are simulated: random deviates are drawn (i) from multivariate normal distributions to reproduce the errors in the parameter estimates, (ii) from univariate normal distributions to account for the tree and growth interval random effects as well as the residual errors in the diameter increment submodel, and (iii) from univariate uniform distributions to determine whether a particular tree died or survived the growth interval. The variances of the multivariate and univariate distributions are estimated during the submodel fitting stage.

For a given Monte Carlo realization, the errors in the parameter estimates are kept constant. A single tree random effect is simulated for each tree. In the mortality submodel, an interval random effect is drawn before each growth interval. All model components were implemented with stochastic variability, except Honer et al.'s total volume equations for which no information on the variance was available in the original publication.

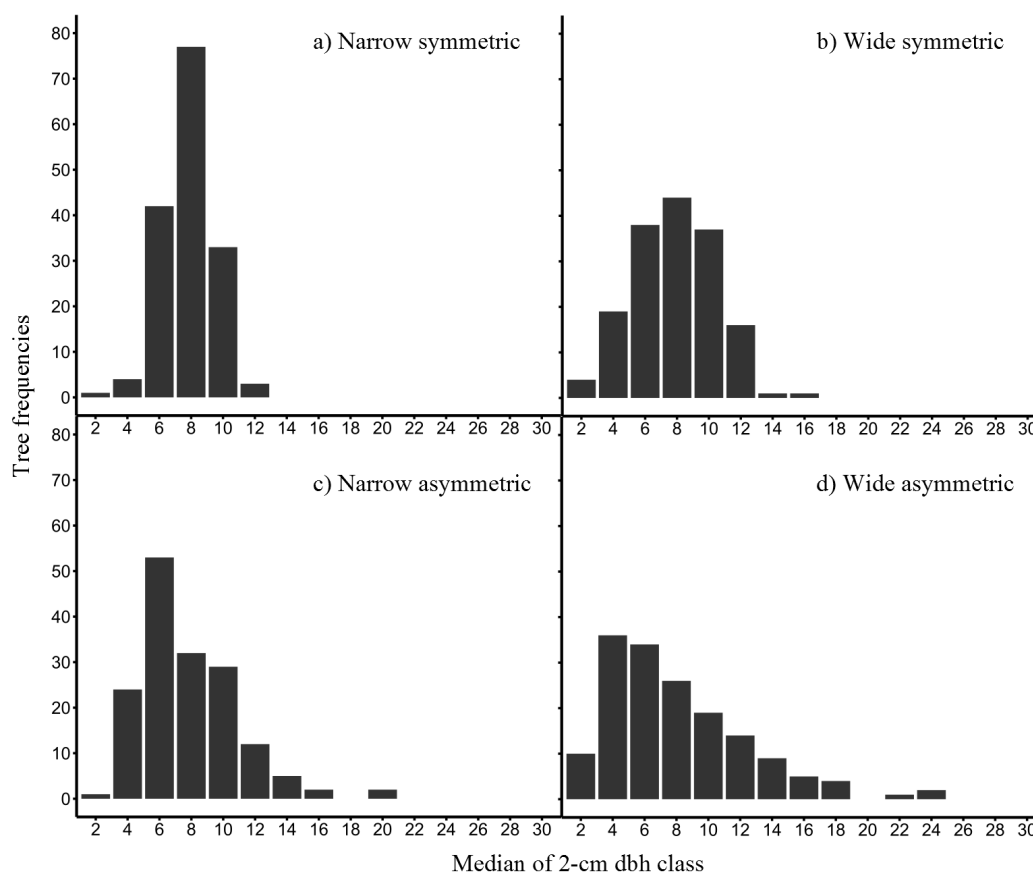


Figure 1. Diameter distributions of the 4000-trees·ha⁻¹ plots in the narrow symmetric, wide symmetric, narrow asymmetric and wide asymmetric sets of plots: normal (a,b) and log normal (c,d).

To test the behaviour of the growth model with respect to the self-thinning relationship, we first created four sets of four 400-m² plots with densities of 4000, 6000, 8000, and 10 000 trees·ha⁻¹. The four sets differed in the way tree dbh were generated in order to test the impact of the diameter distribution on the evolution of each plot. For the first set, the dbh were randomly drawn from a normal distribution with a mean of 8 cm and a standard deviation of 1.5 cm. In the second set, the dbh were also generated from a normal distribution with a mean of 8 cm, but the standard deviation was extended to 2.5 cm. In the third set, the dbh were drawn from a log normal distribution with a mean of 2 and a standard deviation of 0.40 on the log-transformed scale. In the fourth set, the dbh were also generated using a log normal distribution with a mean of 1.95 and a standard deviation of 0.51. These values on the log-transformed scale were selected in such a way that the means on the untransformed scale were also close to 8 cm and therefore comparable to the first two sets. For the sake of clarity, we will refer to these four sets as the narrow

symmetric, the wide symmetric, the narrow asymmetric and the wide asymmetric sets. The diameter distribution of the 4000-trees·ha⁻¹ plot in each set is shown in Fig. 1. For the simulations, we assumed that no spruce budworm defoliation would occur, i.e., $sbw_{ik-1} = 0$ and $sbw_{ik} = 0$.

Once the four sets were created, we ran a 500-realization stochastic simulation over a 70-year period for each plot. At each 5-year growth step, the total volume of each tree was predicted. These volumes were summed up at the plot level and then divided by the number of stems in order to obtain the mean tree total volume. The mean tree total volume was plotted against the stem density for each Monte Carlo realization and compared with Bégin et al.'s (2001) maximum size-density relationship for consistency.

Table 3. Maximum likelihood parameter estimates of the mortality submodel.

Parameter	Effect	Estimate and standard error
β_0	Model intercept	-12.6145 (1.2091)
$\beta_{1,1}$	$dbh_{ijk} \times \text{Other species}$	-0.4397 (0.0456)
$\beta_{1,2}$	$dbh_{ijk} \times \text{Betula spp.}$	-0.4887 (0.0431)
$\beta_{1,3}$	$dbh_{ijk} \times \text{Picea spp.}$	-0.5177 (0.0285)
$\beta_{1,4}$	$dbh_{ijk} \times \text{Balsam fir}$	-0.5477 (0.0147)
$\beta_{2,1}$	$dbh_{ijk} \times sbw_{ik-1} \times \text{Other species}$	n/a
$\beta_{2,2}$	$dbh_{ijk} \times sbw_{ik-1} \times \text{Betula spp.}$	n/a
$\beta_{2,3}$	$dbh_{ijk} \times sbw_{ik-1} \times \text{Picea spp.}$	0.9570×10^{-1} (0.2898×10^{-1})
$\beta_{2,4}$	$dbh_{ijk} \times sbw_{ik-1} \times \text{Balsam fir}$	2.6506×10^{-1} (0.1087×10^{-1})
β_3	dbh_{ijk}^2	1.4955×10^{-2} (0.0675×10^{-2})
$\beta_{4,1}$	$bal_{ijk} \times \text{Other species}$	1.6443×10^{-2} (0.8656×10^{-2})
$\beta_{4,2}$	$bal_{ijk} \times \text{Betula spp.}$	1.6753×10^{-2} (0.7919×10^{-2})
$\beta_{4,3}$	$bal_{ijk} \times \text{Picea spp.}$	3.7944×10^{-2} (0.6040×10^{-2})
$\beta_{4,4}$	$bal_{ijk} \times \text{Balsam fir}$	6.9255×10^{-2} (0.3782×10^{-2})
β_5	$\ln(\Delta t_{ik})$	7.3066 (0.7265)
σ_{int}^2	Variance of the growth interval random effect	0.6869 (0.0925)

3 RESULTS

3.1 SUBMODEL FIT

The AUC value of the mortality submodel was 0.868, which indicated a relatively good accuracy. The bias resulting from the cross-validation of the diameter increment submodel was estimated to be $+0.04 \text{ mm}\cdot\text{yr}^{-1}$, which corresponded to a relative bias of 3.6%. The maximum and restricted maximum likelihood estimates of the parameters of both submodels are shown in Table 3 and Table 4.

Table 4. Restricted maximum likelihood parameter estimates of the dbh increment submodel.

Parameter	Effect	Estimate and standard error
γ_0	Model intercept	1.8390 (0.0439)
$\gamma_{1,1}$	Other species	-0.4383 (0.3400)
$\gamma_{1,2}$	<i>Betula</i> spp.	-0.8929 (0.1961)
$\gamma_{1,3}$	<i>Picea</i> spp.	1.4867 (0.1350)
$\gamma_{1,4}$	Balsam fir	n/a
γ_2	dbh_{ijk}	0.1116 (0.0057)
γ_3	dbh_{ijk}^2	-4.2268×10^{-3} (0.2437×10^{-3})
$\gamma_{4,1}$	$sbw_{ik} \times$ Other species	0.5324 (0.1306)
$\gamma_{4,2}$	$sbw_{ik} \times$ <i>Betula</i> spp.	0.6123 (0.0779)
$\gamma_{4,3}$	$sbw_{ik} \times$ <i>Picea</i> spp.	-0.4542 (0.0794)
$\gamma_{4,4}$	$sbw_{ik} \times$ Balsam fir	-0.4053 (0.0096)
$\gamma_{5,1}$	$bal_{ijk} \times$ Other species	-3.0270×10^{-2} (0.8824×10^{-2})
$\gamma_{5,2}$	$bal_{ijk} \times$ <i>Betula</i> spp.	-2.6094×10^{-2} (0.5316×10^{-2})
$\gamma_{5,3}$	$bal_{ijk} \times$ <i>Picea</i> spp.	-7.1038×10^{-2} (0.4133×10^{-2})
$\gamma_{5,4}$	$bal_{ijk} \times$ Balsam fir	-4.4480×10^{-2} (0.0692×10^{-2})
σ_{tree}^2	Variance of the tree random effect	1.4822×10^{-3} (0.0482×10^{-3})
σ_{res}^2	Variance of the residual error	4.3540×10^{-1} (0.0429×10^{-1})

Mean predicted probabilities of mortality and mean predicted annual dbh increments as a function of tree dbh and the basal area of trees larger than the subject can be found in Fig. 2. The mortality patterns were u-shaped with the lowest predicted probabilities of mortality between 10 and 20 cm in dbh for all species (Fig. 2a).

On average, the other species had higher probabilities of mortality in the absence of spruce budworm defoliation. Increases of the competition index bal_{ijk} resulted in higher mortality probabilities (Fig. 2b). The diameter increment predictions followed the opposite pattern. They reached a maximum of between 10 and 20 cm in dbh (Fig. 2c), and the increase of the bal_{ijk} variable induced a decrease in the predicted increments (Fig. 2d). *Picea* spp. had the largest diameter increments on average.

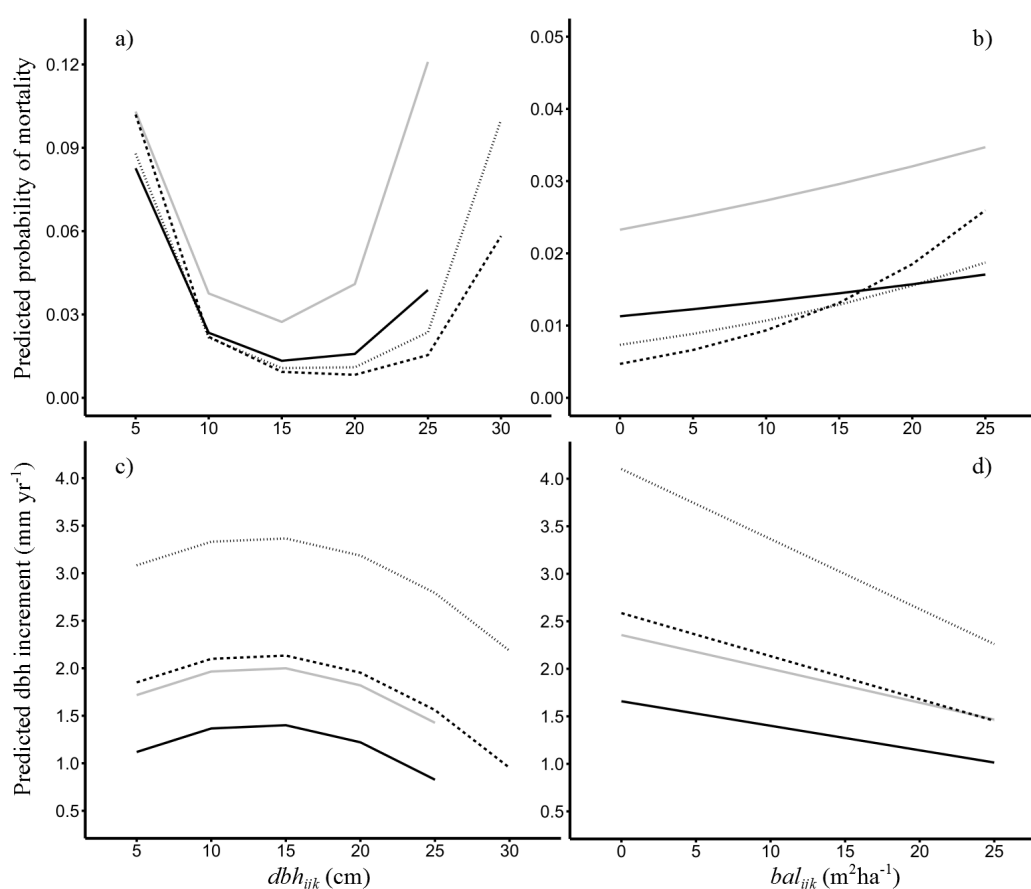


Figure 2. Predicted probabilities of mortality and annual dbh increments as a function of tree dbh (dbh_{ijk}) and basal area for trees whose dbh is larger than the subject (bal_{ijk}) (gray line: Other species, black line: *Betula* spp., gray dotted line: *Picea* spp., black dashed line: Balsam fir; bal_{ijk} is set to $10 \text{ m}^2 \text{ ha}^{-1}$ in a and c; dbh_{ijk} is set to 15 cm in b and d; no spruce budworm outbreak in current and previous growth intervals).

3.2 MODEL FIT ON SINGLE GROWTH INTERVALS

The estimated biases in basal area and stem density predictions are shown in Table 5. Overall, the estimated bias in basal area was less than 1%, whereas the

one in stem density was slightly larger, with a relative value of 3.2%. The first four intervals were those with the largest biases, with relative values as large as -17.1%. The last three intervals had lower relative biases ranging from -1.4% to 7.5%. The biases in stem density tended to be relatively larger than those in basal area. It also appeared that the model tended to overestimate growth during the spruce budworm outbreak while underestimating it during the other intervals.

Table 5. Estimated biases of basal area and stem density predictions on single growth intervals (relative biases appear in parentheses).

Growth interval	Number of observations	Basal area (m ² ha ⁻¹)	Stem density (trees·ha ⁻¹)
1968-1973	32	3.1 (7.4%)	933 (14.9%)
1973-1978	32	2.3 (5.2%)	241 (4.3%)
1978-1983 ^a	32	-5.3 (-16.2%)	-497 (-17.1%)
1983-1989 ^a	32	-1.8 (-6.5%)	-144 (-7.8%)
1989-1994 ^b	23	0.6 (1.4%)	-33 (-1.4%)
1994-2000 ^c	23	1.6 (4.0%)	140 (7.5%)
2000-2006	23	2.8 (6.9%)	100 (6.1%)
All intervals	197	0.3 (0.8%)	106 (3.2%)

^a Growth intervals with spruce budworm outbreak

^b Prior to the 1994 shelterwood cutting

^c Following the shelterwood cutting

3.3 EVALUATION OF LONG-TERM PROJECTIONS AND DIAMETER DISTRIBUTION EFFECT

The stochastic simulations of the narrow symmetric set are shown in Fig. 3. Most Monte Carlo realizations were below the MSDR, except for some of the highest initial density (Fig. 3b,c,d). For all plots, the mean total tree volume tended to increase while the stem density was decreasing.

The decrease in stem density was dependent on the initial densities. For the two plots with the lowest initial stem densities (Fig. 3a,b), the mean tree volume increased more rapidly than the stem densities decreased, dragging the different realizations closer to the MSDR. For the plots with the highest initial stem densities (Fig. 3c,d), the realizations tended to remain close to the MSDR.

An interesting feature that was observed in the simulations of each plot was the

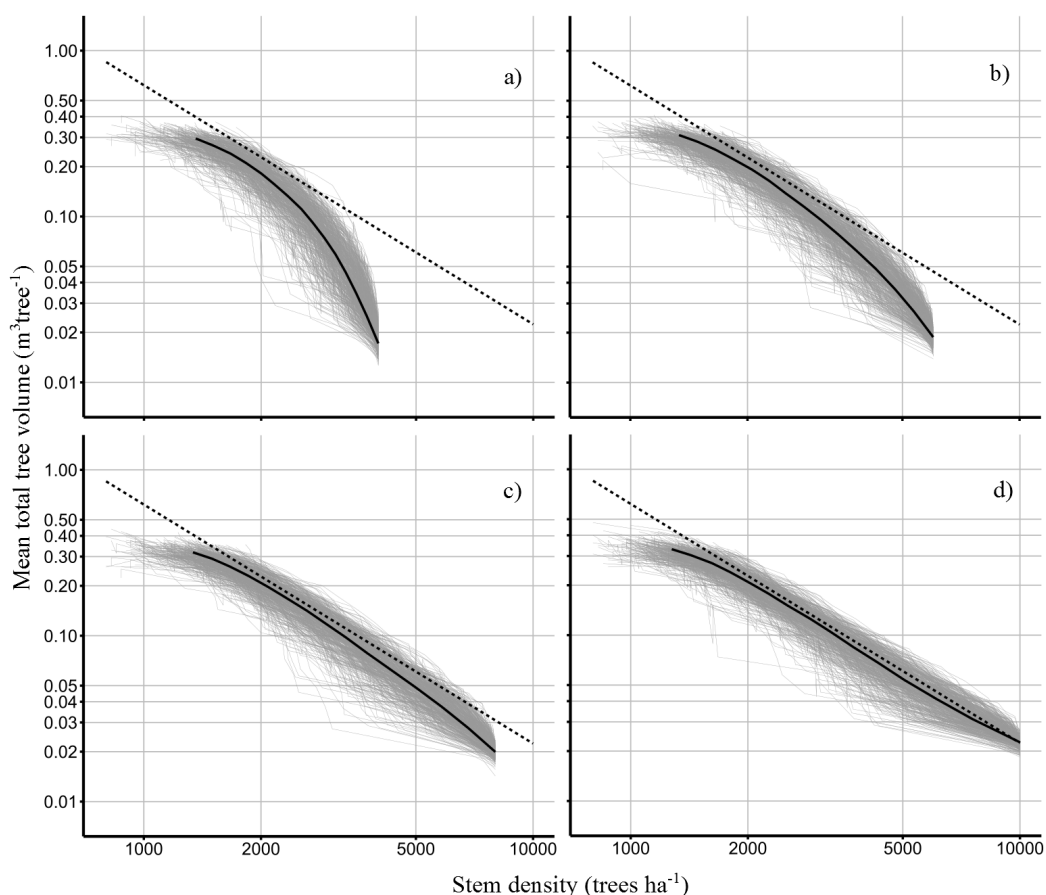


Figure 3. Stochastic simulations for four 400-m² balsam fir plots with initial narrow symmetric dbh distributions and different initial stem densities: 4000 (a), 6000 (b), 8000 (c) and 10 000 (d) trees·ha⁻¹ (simulations based on 500 Monte Carlo realizations, solid black line: mean trend, black dashed line: Bégin et al.'s (2001) maximum size-density relationship).

break up at values between 0.2 and 0.3 m³tree⁻¹. After this point, the total volume increment appeared to slow down while the decrease in stem density maintained the same pace. For all four plots, this slowing down induced a divergence from the MSDR. The simulations with the three other sets showed similar trends.

The mean predicted total tree volumes as a function of the stem density are compared across the different initial distributions in Fig. 4. The initial asymmetric dbh distributions resulted in a greater mean total tree volume at the beginning of the simulations. However, the plots with initial asymmetric dbh distributions tended to show trajectories that flattened out when compared with those of the symmetric distributions. Except for the lowest initial density (Fig. 4a), the effect of the variance was quite small.

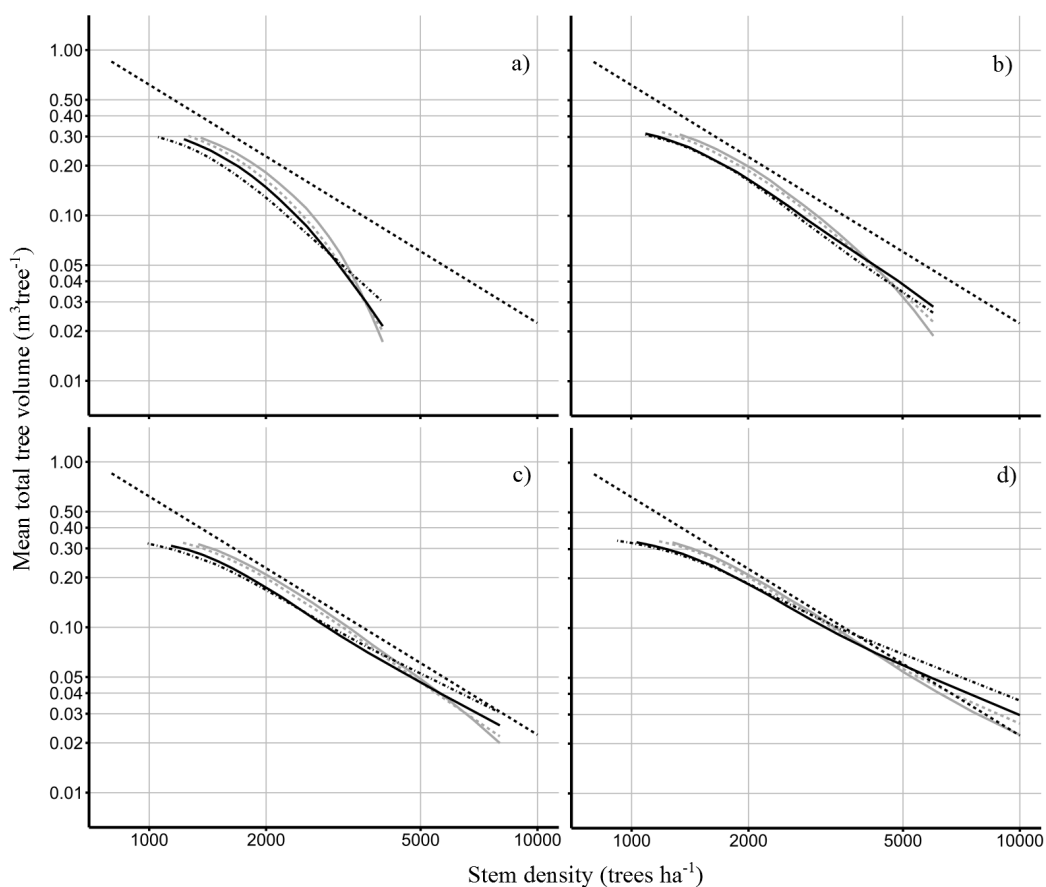


Figure 4. Mean predictions of the mean total tree volume as a function of stem density for four sets of plots with different diameter distributions and different initial stem densities: 4000 (a), 6000 (b), 8000 (c) and 10,000 (d) trees·ha⁻¹ (gray line: narrow symmetric, gray dashed line: wide symmetric, black line: narrow asymmetric, black dotted line: wide asymmetric, black dashed line: Bégin et al.'s (2001) maximum size-density relationship).

4 DISCUSSION

The Matapedia single tree-based model has a simple architecture: it consists of two dynamic components and two static ones. Its dynamic components are a diameter increment submodel and a mortality submodel. These submodels rely on the same predictors, namely tree dbh and species group, basal area of all trees with dbh larger than the subject, and the occurrence of spruce budworm defoliation in the previous or the upcoming growth intervals. Coupled with the two other static components, a height-diameter relationship and Honer et al.'s total tree volume equation, the model makes it possible to run stochastic growth simulations for even-aged balsam fir-dominated stands.

Although the individual components of the simulator were independently evaluated, there was no guarantee that the plot-level predictions would be consistent. The

evaluation over single growth intervals showed that the biases in total basal area and stem density were relatively small on average (Table 5). Although there was some evidence that the model might overestimate the growth during spruce budworm outbreaks and vice versa, the comparison between long-term projections and Bégin et al.'s MSDR yielded surprisingly good results, as shown in Fig. 3.

These results also showed that the relative density indices are not required to make consistent predictions with single tree-based models. Actually, using mainly tree-level variables, we succeeded in reproducing growth trajectories that comply with the self-thinning law. Monserud et al. (2004) conducted a similar study using the PROGNAUS model with pure even-aged stands of Norway spruce (*Picea abies* L. Karst.) and Scots pine (*Pinus sylvestris* L.). They concluded that a density-dependent individual-tree mortality submodel may be sufficient to reproduce the stand-level behaviour inherent in Reineke's (1933) rule. Although we did not use Reineke's index, our findings are in accordance with their conclusion regarding the ability of a diameter increment and a mortality submodel to implicitly reproduce trajectories that closely match Bégin et al.'s (2001) MSDR for pure balsam fir stands (Fig. 3).

The originality of our study is due to the fact that our simulations were stochastic and not deterministic. If the two submodels properly represent the diameter increment and mortality processes in the population, our stochastic simulations then reproduce the variability that might be expected in this population. Bégin et al.'s (2001) modelled the MSDR of pure balsam fir stands in such a way that approximately 95% of the observations would be below the line. Considering that the simulations included 500 realizations and that only some exceeded Bégin et al.'s (2001) MSDR (Fig. 3b,c,d), our results seem to be consistent with their relationship, even though the height-diameter relationship and the total volume equation we used were different from theirs. Further simulations should be run in order to identify the eventual differences between the relationship that is implicit in our simulator and Bégin et al.'s MSDR. It would also be interesting to derive an analytic expression of the MSDR from the submodel specifications. The developments described in Section 2.2.3 might eventually be used for that purpose.

In many growth models, fertility is taken into account through a proxy, such as site index, or some climate or site variables (e.g. Vanclay 1989; Ung et al. 2009). The Matapedia growth model does not rely on any of these variables. In our preliminary attempts to fit the model components, we specified a plot random effect in both submodels. A plot random effect is actually an unobserved factor that has a

systematic and sustained effect on a particular plot (Gregoire 1987). In this context, we expected that this random effect would account for the variability in fertility. It turned out that the plot random effect was not significant in neither of the submodels. We then tried to fit the submodels with climate variables, such as the mean annual temperature and precipitation for the 1971-2000 period, but without any further improvement. The homogeneity of the conditions may explain these unsuccessful attempts in taking fertility into account. All our plots belong to the same forest type and they are geographically close to each other so that the climate and soil conditions are similar. In such conditions, Lafèche et al. (2013, p.30) estimated a site index of 16.6 m at a reference age of 50 years. In this context, the scope of the model remains limited to even-aged balsam fir-dominated stands growing on undifferentiated tills in the balsam fir-yellow birch bioclimatic domain in Eastern Québec.

In these conditions, the Matapedia growth model makes it possible to predict tree and stand development for dense balsam fir stands that are at least 33 years old (Pothier 1998). Compared to previous studies, our growth trajectories for low density stands showed mortality at younger ages even if the stand was still far from the MSDR (Fig. 3a). Many authors have identified an “imminent competition-induced mortality zone” as soon as the relative density (density/maximum density for a given mean tree volume) reaches 0.5-0.6 (Drew and Flewelling 1979; Newton and Weetman 1993; Penner et al. 2006). There are two explanations for the fact that mortality occurred below that threshold in our simulations. Firstly, since those stands have experienced a spruce budworm outbreak, the model may include some losses even though pesticides were used to limit the damage. Secondly, the mortality submodel accounts for all sources of mortality, including the catastrophic mortality due to factors other than spruce budworm. Some windfalls were observed in the dataset, and the mortality submodel reproduces these in the simulations. In that sense, the stochastic simulation includes not only the competition-induced mortality but the catastrophic mortality as well.

The tendency of the trajectories to become progressively parallel to the MSDR as the stands grow is in accordance with the stand density management diagram (SDMD) concepts (cf. Jack and Long 1996). These evolutions are also coherent with the observed development of balsam stands reported by Penner et al. (2006). However, the divergence of growth trajectories with the MSDR at older ages can hardly be predicted with the SDMD alone. This breakup trend that was observed in all our simulations had also been observed in previous studies. VanderSchaaf and

Burkhart (2008) defined three phases in the density-dependent mortality stage of stand development. Phase I occurs when the stand enters the imminent competition mortality zone. The growth trajectory then exhibits a concave shape on the log-log scale as it becomes parallel to the MSDR. In phase II, the stand trajectory is linear since the increase in average tree size is compensated for by a decrease in the number of individuals. Phase III occurs when residual trees can no longer occupy the canopy gaps. This causes a divergence from the MSDR. The breakup we observed in this study is clearly the beginning of the third phase, which is also known as the senescence period.

In trembling aspen even-aged stands, Pothier et al. (2004) managed to determine that the senescence period begins around 60 years of age. Our model is not age dependent but rather size dependent and our simulations were not set to start a particular age. If we assume that the trees were initially the same age than those at Lac Matapedia in 1968, the senescence would begin when the stand reaches 80 years of age in the absence of silvicultural treatment and spruce budworm outbreak. However, the value is highly hypothetical since the recurrence interval of spruce budworm outbreaks was around 30 years during the 20th century (Blais 1983).

In terms of size, the senescence phase begins when the mean total tree volume reaches a value of between 0.2 and 0.3 m³tree⁻¹ in the simulations of the narrow symmetric plot set. According to Honer et al.'s total volume equation, the mean total tree volume of 0.2 m³tree⁻¹ is obtained with a balsam fir tree of 19 cm in dbh and 16 m in height. The value of 0.3 m³tree⁻¹ is obtained with a tree of approximately 23 cm and 18 m in dbh and height, respectively. When trees become larger than 25 cm in dbh, the predicted probability of mortality increases, whereas the annual dbh increment decreases (Fig. 2a,c). This value is a little over 20 cm, considered by McLintock (1954) to be at risk of wind damage. The decrease in the dbh increment is not of concern in the sense that smaller dbh increments on larger trees may still result in similar or even larger basal areas and volume increments than those of small trees with larger dbh increments. On the other hand, the increasing probability of mortality with dbh for a given competition level (Fig. 2a) has a much more significant impact: as mortality losses increase and the volume increment remains the same on survivor trees, it can only be expected that the plot trajectory will diverge from the MSDR.

In their comparison of different crown radius models for balsam fir, Gilmore and Seymour (1997) obtained predicted crown radii as large as 2.0 m for open-grown and co-dominant balsam fir trees. Assuming that the trees would be uniformly

distributed over the plot, nearly 700-800 trees per hectare could fully occupy the canopy. Of course, trees are not evenly distributed, nor do they immediately occupy a gap as soon as it appears. In our simulations, the gaps caused by the increasing mortality appeared at a faster pace than they could be occupied by survivor trees. As the mortality increases, this pace is expected to increase as well and the stand is likely to reach the old-growth stage (Oliver 1981).

The causes for this increasing mortality in trees with larger diameters are exogenous to the between-tree competition. Balsam fir is known to be vulnerable to butt and root rots as well as red heart fungus (*Haematostereum sanguinolentum*). Patry et al. (2004) reported a 50% probability of observing butt rot for 65-year-old balsam fir trees. This age corresponds to the one when root decay becomes a risk (Whitney 1989) and is a little less than the age identified by Basham (1991) at which balsam fir should be harvested in order to avoid serious decay losses. These damaging agents are known to affect the mechanical stability of balsam trees, making them more vulnerable to windthrow as they become older (Frank 1990).

Our simulations on the structure effects show that asymmetric initial diameter distributions resulted in higher mean total tree volume at the beginning of the simulation. This can be explained by the fact that volume is a three-dimensional measurement, whereas diameter is only a single dimension. For the same mean diameter, larger trees account for much more volume than small trees do. In this study, the plots with asymmetric distributions contained more trees larger than 14 cm than the plots with symmetric distributions (Fig. 1). The estimated volumes of those trees pull the mean total tree volume upwards for the same density, which explains the higher initial mean total tree volume for the plot of the asymmetric sets (Fig. 4).

On the other hand, the simulations of the plots in these asymmetric sets showed that their mean total tree volume was lower on average at advanced ages (Fig. 4). The explanation for this result seems to be related to the above-mentioned arguments. In fact, the plots of the asymmetric sets have a typical right-tailed dbh distribution (Fig. 1). Individually, these large trees reach the 25-cm threshold in dbh sooner than those in the plots of the symmetric sets. When they reach this threshold, these trees are more prone to die, which results in higher losses in mean total tree volume than in the plots of the symmetric sets. The MSDRs that flatten out in the asymmetric sets are essentially due to different individual dynamics.

Monserud et al. (2004) compared PROGNAUS predictions with two existing MSDRs: Reineke's (1933) relationship and Sterba's (1987) index. When compared to Sterba's MSDR, the maximum site density indices in the simulations were much

lower. At least a part of this difference can be explained by methodological issues. Monserud et al. (2004) ran deterministic simulations. However, when they used stochastic simulations, they only considered the means of these simulations. An individual-based model represents the mathematical expectation of the growth process. A high degree of uncertainty still exists around model predictions (Kangas 1999; Fortin et al. 2009a).

On the other hand, MSDRs are usually fitted on plot data, which are realizations of the growth process. The plots that define the MSDR are those with the highest stem densities for a given size. Those plots are the highest percentiles in all the realizations of the growth process. Unless the prediction error variance is small, these realizations can be far from the average trend by definition. This difference can be found in Fig. 3. We recall that Bégin et al.'s MSDR was fitted in such a way that 95% of the observations were under the maximum size-density line. If we had only considered the mean predicted values (solid black line), we would have concluded that the maximum size-density relationship in our model was slightly lower than the one in Bégin et al. Taking the uncertainty associated with the predictions into account, as we did in our Monte Carlo simulations, is clearly more in accordance with the way the MSDR are usually fitted.

The prediction uncertainty is also closely related to the plot size. In this study, we were faced with the ideal situation: the plots Bégin et al. used to fit their MSDR and the plots that served to fit the model had the same area, namely 400 m² each. All other things considered, repeating this study with a model whose components would be fit with smaller or larger plots is likely to result in something different. The variance in basal area would probably increase as the plot size decreases. If no bias is to be expected here, the variance of the predictions would definitely be different. In other words, with the submodels fitted on smaller plots, we would expect many more realizations to exceed the MSDR and vice versa for larger plots.

5 CONCLUSIONS

A simple mortality submodel coupled with a diameter increment submodel may implicitly reproduce existing MSDRs. We do not suggest this is always the case. However, in our case study, the adequacy between Monte Carlo-simulated evolutions and the MSDR was surprisingly good, considering that the predictors in our submodels were only the species group, tree dbh and the basal area of all trees with a dbh larger than the subject. Furthermore, trajectories are also consistent with the conceptual or the observed development of balsam fir stands on SDMD.

The simulations made it possible to identify the threshold beyond which there is a divergence from the MSDR. For balsam fir trees in the Lac Matapedia experiment, the threshold for the dbh could roughly be set to 25 cm. As soon as the trees reach this threshold, their probability of mortality drastically increases.

Skewed diameter distributions might result in a change in the slope of the MSDR, with higher mean total tree volumes for the same density, and even more so with greater initial stem densities and vice versa. This impact seems to be closely linked to the aforementioned threshold. In right-skewed diameter distributions, some trees reach the 25-cm dbh threshold sooner than in symmetric distributions for the same mean diameter.

ACKNOWLEDGEMENTS

This project was funded (Grant 112310009) by the Quebec Ministry of Natural Resources. The UMR 1092 LERFoB is supported by the French National Research Agency through the Laboratory of Excellence ARBRE (ANR-12 LABXARBRE-01). The authors are grateful to Jean Bégin, professor at Laval University, and his research team for the 1994 and 2000 measurements at the Lac Matapedia Experimental Forest. The original experimental design was installed by René Doucet, retired scientist at the Forest Research Department of the Quebec Ministry of Natural Resources. Special thanks are also due to the Associate Editor and two anonymous reviewers, as well as Isabelle Auger and Jean-Pierre Saucier from the Quebec Ministry of Natural Resources, for their helpful comments on a preliminary version of this paper.

REFERENCES

- Basham, J. T. (1991). Stem decay in living trees in ontario's forests: a user's compendium and guide. Technical report, Forestry Canada, Ontario Region, Sault Ste. Marie, Ontario.
- Bégin, E., Bégin, J., Bélanger, L., Rivest, L. P., and Tremblay, S. (2001). Balsam fir self-thinning relationship and its constancy among different ecological regions. *Canadian Journal of Forest Research*, 31:950–959.
- Belcher, D. W., Holdaway, M., and Brand, G. (1982). A description of stems. the stand and tree evaluation and modeling system. Technical report, USDA Forest Service.
- Blais, J. R. (1983). Trends in frequency, extent, and severity of spruce budworm outbreaks in eastern canada. *Canadian Journal of Forest Research*, 13:539–547.
- Drew, T. J. and Flewelling, J. W. (1979). Stand density management: an alternative approach and its application to douglas-fir plantations. *Forest Science*, 25:518–532.
- Dufour-Kowalski, S., Courbaud, B., Dreyfus, P., Meredieu, C., and de Coligny, F. (2012). Capsis : an open software framework and community for forest growth modelling. *Annals of Forest Science*, 69:221–233.
- Fontes, L., Bontemps, J. D., Bugmann, H., Van Oijen, M., Gracia, C., Kramer, K., Linder, M., Rötzer, T., and Skovsgaard, J. P. (2010). Models for supporting forest management in a changing environment. *Forest Systems*, 19:8–29.
- Fortin, M. (2013). Population-averaged predictions with generalized linear mixed-effects models in forestry: an estimator based on gauss-hermite quadrature. *Canadian Journal of Forest Research*, 43:129–138.
- Fortin, M., Bédard, S., DeBlois, J., and Meunier, S. (2008). Predicting individual tree mortality in northern hardwood stands under uneven-aged management in southern québec, canada. *Annals of Forest Science*, 65:1–12.
- Fortin, M., Bédard, S., DeBlois, J., and Meunier, S. (2009a). Assessing and testing prediction uncertainty for single tree-based models: A case study applied to northern hardwood stands in southern québec, canada. *Ecological Modelling*, 220:2770–2781.
- Fortin, M., Bernier, S., Saucier, J. P., and Labbé, F. (2009b). Une relation hauteur-diamètre tenant compte de l'influence de la station et du climat pour 20 espèces commerciales du québec. Technical report, Ministère des Ressources naturelles et de la Faune du Québec, Direction de la recherche forestière.

- Frank, R. M. (1990). *Silvics of North America*, chapter Balsam fir, pages 26–35. USDA Forest Service.
- Gilmore, D. W. and Seymour, R. S. (1997). Crown architecture of abies balsamea from four canopy positions. *Tree physiology*, 17:71–80.
- Gregoire, T. G. (1987). Generalized error structure for forestry yield models. *Forest Science*, 33:423–444.
- Gregoire, T. G., Lin, Q. F., Boudreau, J., and Nelson, R. (2008). Regression estimation following the square-root transformation of the response. *Forest Science*, 54:597–606.
- Grondin, P., Gosselin, J., Saucier, J. P., Morneau, C., and Dallaire, S. (2013). *Le guide sylvicole du Québec. Les fondements biologiques de la sylviculture. Tome I.*, chapter La dynamique des peuplements et les végétations potentielles. Les publications du Québec, Québec, Canada.
- Honer, T. G., Ker, M. F., and Alemdag, I. S. (1983). Metric timber tables for the commercial tree species of central and eastern Canada. Technical report, Maritimes Forest Research Centre.
- Jack, S. B. and Long, J. N. (1996). Linkages between silviculture and ecology: an analysis of density management diagrams. *Forest Ecology and Management*, 86:205–220.
- Kangas, A. S. (1999). Methods for assessing uncertainty of growth and yield predictions. *Canadian Journal of Forest Research*, 29:1357–1364.
- Lafèche, V., Bernier, S., Saucier, J. P., and Gagné, C. (2013). Indices de qualité de station des principales essences commerciales en fonction des types écologiques du Québec méridional. Technical report, Gouvernement du Québec, Ministère des Ressources naturelles, Direction des inventaires forestiers.
- Lasko, T. A., Bhagwat, J. G., Zou, K. H., , and Ohno-Machado, L. (2005). The use of receiver operating characteristic curves in biomedical informatics. *Journal of Biomedical Informatics*, 38:404–415.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models, Second Edition*. Chapman and Hall / CRC, New York, USA.
- McLintock, T. F. (1954). Factors affecting wind damage in selectively cut stands of spruce and fir in Maine and northern New Hampshire. Technical report, USDA Forest Service.
- Meyer, W. H. (1929). Yields of second-growth spruce and fir in the northeast. Technical report, USDA Forest Service.

- Monserud, R. A., Ledermann, T., and Sterba, H. (2004). Are self-thinning constraints needed in a tree-specific mortality model? *Forest Science*, 50:848–858.
- Newton, P. F. and Weetman, G. F. (1993). Stand density management diagrams and their development and utility in black spruce management. *The Forestry Chronicle*, 69:421–430.
- Oderwald, R. G. and Hans, R. P. (1993). Corroborating models with model properties. *Forest Ecology and Management*, 62:271–283.
- Oliver, C. D. (1981). Forest development in north america following major disturbances. *Forest Ecology and Management*, 3:153–168.
- Patry, A., Côté, L., Pelletier, D., Riopel, M., and Saucier, J. P. (2004). Analyse de la croissance et de la carie des essences résineuses. régions du bas-saint-laurent et de la gaspésie. Technical report, Gouvernement du Québec, Ministère des Ressources naturelles, de la Faune et des Parcs, Direction des inventaires forestiers.
- Penner, M., Swift, D. E., Gagnon, R., and Brissette, J. (2006). A stand density management diagram for balsam fir in new brunswick. *The Forestry Chronicle*, 82:700–711.
- Porté, A. and Bartelink, H. H. (2002). Modelling mixed forest growth: a review of models for forest management. *Ecological Modelling*, 150:141–188.
- Pothier, D. (1998). Développement de sapinières éclaircies exposées à une épidémie de tordeuse des bourgeons de l'épinette. *The Forestry Chronicle*, 74:91–99.
- Pothier, D., Raulier, F., and Riopel, M. (2004). Ageing and decline of trembling aspen stands in quebec. *Canadian Journal of Forest Research*, 34:1251–1258.
- Pretzsch, H., Biber, P., and Ďurský, J. (2002). The single tree-based stand simulator silva: construction, application and evaluation. *Forest Ecology and Management*, 162:3–21.
- Reineke, L. H. (1933). Perfecting a stand-density index for even-aged forests. *Journal of Agricultural Research*, 46:627–638.
- Robitaille, A. and Saucier, J. P. (1998). *Paysages régionaux du Québec méridional*. Gouvernement du Québec, Ministère des Ressources naturelles. Direction des inventaires forestiers, Québec, Canada.
- Stage, A. R. (1973). Prognosis model for stand development. Technical report, USDA Forest Service.
- Sterba, H. (1987). Estimating potential density from thinning experiments and inventory data. *Forest Science*, 33:1022–1034.

- Thibault, D., Bégin, J., and Bélanger, L. (1995). Relations entre les indicateurs de croissance du sapin baumier en début d'épidémie et sa vulnérabilité à la tordeuse des bourgeons de l'épinette. *Canadian Journal of Forest Research*, 25:1292–1302.
- Ung, C. H., Bernier, P. Y., Guo, X. J., and Lambert, M. C. (2009). A simple growth and yield model for assessing changes in standing volume across canada's forests. *The Forestry Chronicle*, 85:57–64.
- Vanclay, J. K. (1989). A growth model for north queensland rainforests. *Forest Ecology and Management*, 27:245–271.
- Vanclay, J. K. (1994). *Modelling forest growth and yield. Applications to mixed tropical forests*. CAB International, Wallingford, UK.
- Vanclay, J. K. and Skovsgaard, J. P. (1997). Evaluating forest growth models. *Ecological Modelling*, 98:1–12.
- VanderSchaaf, C. L. and Burkhart, H. E. (2008). Using segmented regression to estimate stages and phases of stand development. *Forest Science*, 54:167–175.
- Whitney, R. D. (1989). Root rot damage in regenerated stands of spruce and balsam fir in ontario. *Canadian Journal of Forest Research*, 19:295–308.

A APPENDIX

EXPECTATIONS OF PLOT BASAL AREA AND STEM DENSITY FOR A SINGLE GROWTH INTERVAL

Using the rule of iterated expectations, the expectation of the plot basal area can be expressed as:

$$\begin{aligned}
 E[G_{ik}] &= E[E[G_{ik} | b_{ik}]] \\
 &= E \left[E \left[\sum_j e^{-e^{x_{ijk}\beta + b_{ik}}} \frac{25\pi(10 \cdot dbh_{ijk} + \Delta t_{ik}(\mathbf{w}_{ijk}\boldsymbol{\gamma} + b_{ij}dbh_{ijk} + \varepsilon_{ijk}))^2}{4000000} \mid b_{ik} \right] \right] \\
 &= E \left[\sum_j e^{-e^{x_{ijk}\beta + b_{ik}}} E[g_{ijk}] \right] \tag{A.1}
 \end{aligned}$$

where

$$E[g_{ijk}] = \frac{25\pi}{4000000} \left(\frac{100 \cdot dbh_{ijk}^2 + 20 \cdot dbh_{ijk}\Delta t_{ik}\mathbf{w}_{ijk}\boldsymbol{\gamma} + \Delta t_{ik}^2(\mathbf{w}_{ijk}\boldsymbol{\gamma})^2 + \Delta t_{ik}^2(\sigma_{tree}^2 dbh_{ijk}^2 + \sigma_{res}^2)}{\Delta t_{ik}^2(\mathbf{w}_{ijk}\boldsymbol{\gamma})^2 + \Delta t_{ik}^2(\sigma_{tree}^2 dbh_{ijk}^2 + \sigma_{res}^2)} \right) \tag{A.2}$$

The marginal expectation A.1 is obtained by marginalizing out the random effect b_{ik} . As mentioned in Section 2.2.1, this can be done by integrating the expectation conditional on b_{ik} over the distribution of b_{ik} :

$$\begin{aligned}
 E[G_{ik}] &= E \left[\sum_j e^{-e^{x_{ijk}\beta + b_{ik}}} E[g_{ijk}] \right] \\
 &= \int \left(\sum_j e^{-e^{x_{ijk}\beta + b_{ik}}} E[g_{ijk}] \right) \varphi(b_{ik}; \sigma_{int}^2) db_{ik} \\
 &= \sum_j \int e^{-e^{x_{ijk}\beta + b_{ik}}} E[g_{ijk}] \varphi(b_{ik}; \sigma_{int}^2) db_{ik} \\
 &= \sum_j E[g_{ijk}] \int e^{-e^{x_{ijk}\beta + b_{ik}}} \varphi(b_{ik}; \sigma_{int}^2) db_{ik} \\
 &= \sum_j E[g_{ijk}] E[e^{-e^{x_{ijk}\beta + b_{ik}}}]
 \end{aligned}$$

The term $E[e^{-e^{x_{ijk}\beta + b_{ik}}}] = \int e^{-e^{x_{ijk}\beta + b_{ik}}} \varphi(b_{ik}; \sigma_{int}^2) db_{ik}$ cannot be solved analytically, and numerical integration has to be used instead. Just as in Section 2.2.1, the Gauss-Hermite quadrature was used to approximate this integral.

Following the same developments, the expectation of the stem density is:

$$E[N_{ik}] = \sum_j E[e^{-e^{x_{ijk}\boldsymbol{\beta}+b_{ik}}}]$$

It is worth mentioning that the parameters $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, σ_{tree}^2 , σ_{int}^2 , and σ_{res}^2 are unknown. To obtain feasible estimators for the basal area and the stem density, these parameters are replaced by their maximum or restricted maximum likelihood estimators. The errors in the parameter vector $\hat{\boldsymbol{\gamma}}$ are likely to induce a bias in the square transformation of the dbh in Eq. A.2 as shown by Gregoire et al. (2008). The larger the variance of the errors, the larger the bias is. Considering the extent of our dataset, we assumed that the errors in these parameter estimates were negligible.