Improving tree mortality models by accounting for environmental influences

Henrik Hartmann, Christian Messier, and Marilou Beaudet

Abstract: Tree-ring chronologies have been widely used in studies of tree mortality where variables of recent growth act as an indicator of tree physiological vigour. Comparing recent radial growth of live and dead trees thus allows estimating probabilities of tree mortality. Sampling of mature dead trees usually provides death-year distributions that may span over years or decades. Recent growth of dead trees (prior to death) is then computed during a number of periods, whereas recent growth (prior to sampling) for live trees is computed for identical periods. Because recent growth of live and dead trees is then computed for different periods, external factors such as disturbance or climate may influence growth rates and, thus, mortality probability estimations. To counteract this problem, we propose the truncating of live-growth series to obtain similar frequency distributions of the "last year of growth" for the populations of live and dead trees. In this paper, we use different growth scenarios from several tree species, from several geographic sources, and from trees with different growth patterns to evaluate the impact of truncating on predictor variables and their selection in logistic regression analysis. Also, we assess the ability of the resulting models to accurately predict the status of trees through internal and external validation. Our results suggest that the truncating of live-growth series helps decrease the influence of external factors on growth comparisons. By doing so, it reinforces the growth–vigour link of the mortality model and enhances the model's accuracy as well as its general applicability. Hence, if model parameters are to be integrated in simulation models of greater geographical extent, truncating may be used to increase model robustness.

Résumé : La dendrochronologie été largement utilisée dans les études portant sur la mortalité des arbres où des variables de croissance récente sont utilisées comme indicateur de la vigueur physiologique des arbres. La comparaison de la croissance radiale récente d'arbres vivants et morts permet donc d'estimer la probabilité de mortalité des arbres. L'échantillonnage d'arbres matures morts fournit généralement la distribution des années de mortalité qui peuvent s'étendre sur plusieurs années ou décennies. La croissance récente des arbres morts (avant leur mort) est ensuite calculée pour un certain nombre de périodes alors que celle des arbres vivants (avant leur échantillonnage) est calculée pour des périodes identiques. Puisque la croissance récente des arbres vivants et morts est ensuite calculée pour des périodes différentes, des facteurs externes tels les perturbations ou le climat peuvent influencer le taux de croissance et, par conséquent, l'estimation de la probabilité de mortalité. Pour résoudre ce problème, nous proposons de tronquer les séries de croissance des arbres vivants de façon à obtenir des distributions de fréquence similaires de « la dernière année de croissance » pour les populations d'arbres vivants et morts. Dans cette étude, nous utilisons différents scénarios de croissance à partir de plusieurs espèces d'arbre et de plusieurs origines géographiques ainsi que différents patrons de croissance pour évaluer l'impact des séries tronquées sur les variables de prédiction et sur leur sélection dans les analyses de régression logistique. De plus, nous évaluons la capacité des modèles qui en résultent à prédire avec exactitude le statut des arbres à l'aide d'une validation interne et externe. Nos résultats indiquent que les séries de croissance tronquées des arbres vivants contribuent à diminuer l'influence des facteurs externes sur les comparaisons de croissance. De ce fait, elles renforcent le lien entre la croissance et la vigueur dans le modèle de mortalité et améliorent l'exactitude et l'applicabilité générale du modèle. Par conséquent, si les paramètres du modèle doivent être intégrés dans des modèles de simulation à plus grande portée géographique, les séries tronquées peuvent être utilisées pour augmenter la robustesse du modèle.

[Traduit par la Rédaction]

Introduction

Tree mortality is a critical component of forest dynamics. Since the early 1980s, there has been a marked increase in publications related to tree mortality (e.g., Waring 1987;

H. Hartmann,[1] C. Messier, and M. Beaudet. Département des Sciences biologiques, Centre d'Étude de la Forêt (CEF), Université du Québec à Montréal, C.P. 8888, succursale Centre-ville, Montréal, QC H3C 3P8, Canada.

[1]Corresponding author (e-mail: henrik333@sympatico.ca).

Franklin et al. 1987). During this period, some of the earlier studies focused on predicting individual tree mortality in an empirical manner using stem diameter and diameter increment as predictor variables (e.g., Monserud 1976; Buchman 1983; Buchman and Lentz 1984).

Generally, a tree dies when it cannot acquire or mobilize enough resources to repair damage, overcome stress, or otherwise sustain its life (Waring 1987). There are many potential physiological causes for a tree's decline in vigour (Franklin et al. 1987). Tree vigour is a somewhat ambiguous concept aiming to describe a tree's vitality. Tree vigour can be estimated in the field with a visual assessment of the social position of trees and morphological and pathological qualities of the tree stem, crown, or bark (e.g., Ouellet and

Zarnovican 1988; Millers et al. 1991; OMNR 2004) and also with measures of physiological processes (e.g., photosynthesis) or vital functions such as radial growth (Gehrig 2004). Manion's (1981) described a conceptual tree decline model in which the downward spiral toward death is often triggered by some form of disturbance or by the interacting effects of environmental factors and pathogens. If the underlying physiological processes do not yield sufficient synthate to sustain all essential vital functions, tree vigour declines. Oliver and Larson (1996) provided some insight into this concept by ranking the vital functions of a tree in order of allocation priority, where maintenance of live tissue, fine root production, and reproduction precede height and diameter growth. Because radial growth has a low priority in carbon allocation, it is sensitive to the overall carbon balance of a tree and is considered to be positively correlated with tree vigour (Waring and Pitman 1985; Pedersen 1998a, 1998b). Because tree vigour itself is expected to be negatively correlated with tree mortality, radial growth has been successfully used to predict mortality probabilities (e.g., Ogle et al. 2000; Bigler and Bugmann 2003, 2004a, 2004b).

Therefore, low growth rates in dying trees is information that can be used for estimating mortality probabilities using measures of "recent radial growth" of live and dead trees (Wyckoff and Clark 2000). A tree's last year of growth, or the average growth over some period prior to death, can then be used as a predictor variable in, for example, a logistic regression model (e.g., Flewelling and Monserud 2002; Van Mantgem et al. 2003). To account for the fact that trees with slow but steady growth can survive over long periods, whereas trees with initially rapid but then decreasing growth levels often die, some authors also included growth trend variables as predictors for tree mortality models (Bigler and Bugmann 2003, 2004a).

When using logistic models, the prediction of individual-tree mortality probabilities requires live and dead tree growth series, from which various predictor variables, such as those mentioned above, can be computed. However, since the death of individual trees is a relatively rare event in the absence of severe or large-scale disturbance, samples of dead trees usually comprise trees that have died over a more or less wide range of years. On the other hand, the last year of growth of live individuals usually corresponds to the year of sampling or, if growth has not ceased by the time of sampling, the year prior to sampling. Therefore, the anchor point for predictor variable computations (death year or last year of growth) varies between live and dead individuals. This means that computed growth variables for dead trees correspond to different time windows, whereas growth variables of all live individual are computed for the same period. If growing conditions change through time because of the influence of external factors (e.g., disturbance, climate), the recent growth of live trees might be subjected to the influence of factors that will not necessarily have affected the growth of dead trees prior to their death. Therefore, one can expect that the difference in growth (levels and (or) trends) between live and dead trees might be over- or under-estimated relative to what it would have been if it had been evaluated at corresponding periods (Fig. 1). Therefore, under some circumstances, the resulting estimates of

Fig. 1. Recent radial growth series and computed growth variables for live (solid line) and dead (shaded line) individuals. Lines with open symbols indicate the slope of the last 10 years of growth for dead (○, 1986–1995) and live (△, 1986–1995; □, 1995–2004) trees; solid symbols indicate the 5 year median of "recent growth". Note the change in both median and slope when different time frames are used to compute the growth variables of the live tree.
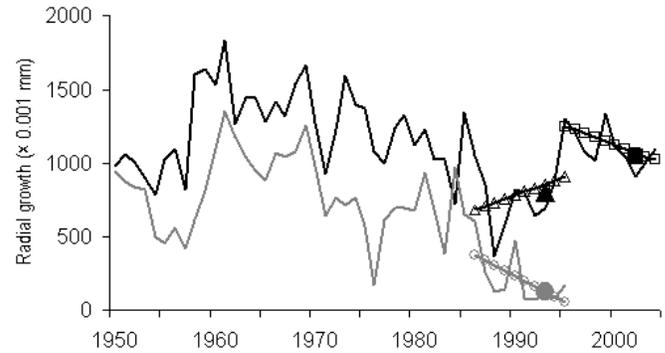
mortality probabilities might be inaccurate. Other researchers have addressed this issue by explicitly modelling environmental variations and intervention occurrences (i.e., inciting stresses) that are common to all trees at a particular site and, then, using only estimated parameters of "vigour-related" growth variations (e.g., Pedersen 1998b). However, by relying completely on model estimates, this method may be prone to add further modelling uncertainty (i.e., through model assumptions and parameter estimate uncertainties) to the resulting mortality model. Truncating constitutes a more direct method and should yield growth variables that reflect more accurately the difference in vigour between live and dead individuals rather than the difference in growing conditions between different time windows. Therefore, the resulting model is expected to better reflect the biological differences occurring between live and dead trees, reducing the need for empirical calibration (Hawkes 2000).

In this paper, our objective was to determine if truncating (right-censoring) the growth series of each live tree to the death year of a paired dead individual would affect the discriminative ability of a logistic individual-tree mortality model. More specifically, we wanted to evaluate the magnitude, if any, of the effect that might be introduced when the proposed procedure (truncating) is not performed and identify situations where it might be especially important to take into account. To do so, we compare models based on truncated live series with those based on untruncated data using predictor choice and validation measures as evaluation criteria.

The overall hypothesis is that the impact of truncating depends on the underlying growth dynamics of the training and validation data sets (where the training data is used to parameterize the model, whereas the validation data is an independent data set used to test the model performance). Two types of validation will be performed: internal and external. Internal validation uses a resample (e.g., bootstrap sample) of the original (training) data, whereas external validation is based on a completely independent data set that does not come from the same sample population (Harrell 2001). We evaluate the model on its discriminative ability, i.e., its ability to correctly predict the status (live or dead) of the trees

from their growth history. We predict that (*i*) an untruncated model should generally have a higher discriminative and predictive ability in internal validation if, during the years when their growth is being compared, the growth of live and dead trees shows a diverging trend over time. Here, truncating would reduce the values of growth level variables of live trees and, therefore, also reduce the differential between live and growth trees. When live and dead tree growth series approach over time (merging trend), (*ii*) truncating should increase the internal predictive ability of the model, because it causes an increase in growth level differences. Because we wish our testing to be as stringent as possible, we also use external validation. Here, (*iii*) truncated models should have a better discriminative ability in most cases. Because truncating possibly eliminates growth variation due to data-specific external factors (e.g., disturbance or climate), it is expected to make these models more indicative of the biological processes of vigour decline preceding death which is the overall assumption of all mortality models based on radial growth.

Materials and methods

Data sources

The three hypotheses were tested using five data sets corresponding to various tree species from different geographic areas: (A) white spruce (*Picea glauca* (Moench) Voss) from the Abitibi region in northwestern Quebec, Canada (48°30′N, 79°1′W) (Senecal et al. 2004), (B) sugar maple (*Acer saccharum* Marsh.) from the Temiscaming area (46°43′N, 79°04′W) (H. Hartmann, unpublished data), (C) balsam fir (*Abies balsamifera* (L.) P. Mill.), (D) black spruce (*Picea mariana* (P. Mill.) BSP) from the lower north shore of the St. Lawrence estuary in eastern Quebec (49°36′N, 68°39′W) (L. DeGrandpré, L., unpublished data), and (E) Norway spruce (*Picea abies* (L.) Karst.) from Pallas Yllas Tunturi in northwestern Finland (67°56′N, 23°44′E) (M.-N.Caron, unpublished data). These stands were of different age structures (uneven aged and even aged) and of different ages (approximately 70–150 years). The stands represented by these data sets underwent different disturbances (e.g., defoliation or drought), which produced different growth dynamics (Fig. 2). In data sets A and C, the growth of live and dead trees diverges over time (at least if one considers the last 15–20 years), whereas live and dead tree growth series in data sets D and E show a merging or non-diverging trend. As for data set B, the mean growth rate of dead trees is lower than that of live trees before 1978 (corresponding to a spruce budworm, *Choristoneura fumiferana* Clem., epidemic) but shows a marked release afterwards, leading to a crossover of live and dead tree growth series. The mean number of years since death for dead trees varied from 3.9 years (data set A) to 22.9 years (data set E) (Table 1).

Data treatment

Among the sampled trees available in each data set, we only used trees with diameters between 19 cm and 49 cm in DBH (diameter at breast height) to avoid heavily suppressed (smaller diameter) and senescent (larger diameter) trees. Trees were selected so that the sample was evenly distrib-

uted within these diameter limits. Live trees were sampled using an increment borer at DBH, whereas cross sections of dead trees were taken at the same height. Radial increment was measured using a computer-assisted micrometer (0.001 mm precision) equipped with a microscope. Live tree growth series were used to construct a master chronology using COFECHA (Holmes 1983) and visual examination of marker years which permitted crossdating of dead individuals. Some of the growth series could not be crossdated with absolute certainty so they were excluded from our analyses. These series were mostly from heavily suppressed trees with very low growth rates and little growth variation. However, the remaining trees showed clear evidence for some kind of cyclic disturbance (i.e., spruce budworm in data sets A, C, and D or forest tent caterpillar, *Malacosoma disstria* Hübner, in data set B) or very cold summer temperatures (data set E), and the associated growth declines served as reliable marker years. Based on these visual datings, COFECHA was used to detect missing or false rings, which would then be identified on the cores or cross sections. After adding or removing these measurements, COFECHA was used again to verify the cross-dating, which usually yielded satisfactorily results.

The original data sets had varying sample sizes, and in most cases, at least two increment cores (live) or two radii on cross sections (dead) were available per tree. These tree-level measurements were averaged to account for intratree variability of radial increment due to growing conditions or leaning (Kienholz 1930; Peterson and Peterson 1995). However, all the tests were run on data sets with an equal number of live and dead individuals to minimize the influence of sample size on predictive ability (Fielding and Bell 1997). To do so, we randomly selected 30 dead and 30 live individuals from each data set. Because the goal of this study was not to estimate absolute mortality probabilities but rather to compare changes in discriminative ability induced by truncating within each data set, we assume that the differences in sampling strategies (e.g., coring height, use of cores vs. cross sections) do not affect the general conclusions.

Thirty live and 30 dead trees were randomly selected form each data set and paired using tree size classes as grouping factor (i.e., DBH classes: 19.1–29.0 cm, 29.1–39.0 cm, and 39.1–49.0 cm) when information on tree size was available (sugar maple, black spruce, and balsam fir). The use of these coarse classes controls for diameter-related growth differences and reduces the risk that live trees, after truncating and therefore reducing diameters, were paired with consistently larger dead trees. The model comparisons should still be valid even where tree size distributions could not be homogenized across live and dead trees, because the same data set was used to compare the truncated with the untruncated model.

Truncating was done by snipping the growth series of a live tree at the year of death of its paired dead individual. Each of the resulting data sets contained 30 dead tree growth series and 30 live tree growth series, the latter as both truncated and untruncated series. A number of variables were computed to describe the recent growth history for each growth series: growth level variables included the medians over 3, 5, and 10 years of "recent" growth (Med_3, Med_5, and Med_10, respectively) and growth-trend varia-

Fig. 2. Mean chronologies for live (solid line) and dead (shaded line) tree growth series (1950 to the end of the series) and sample size of dead individual (broken line) for data sets A to E. Thirty live trees were sampled throughout the period. Figure lettering corresponds to data set identification for the different data sets (A–E); see Table 1 and text for details.
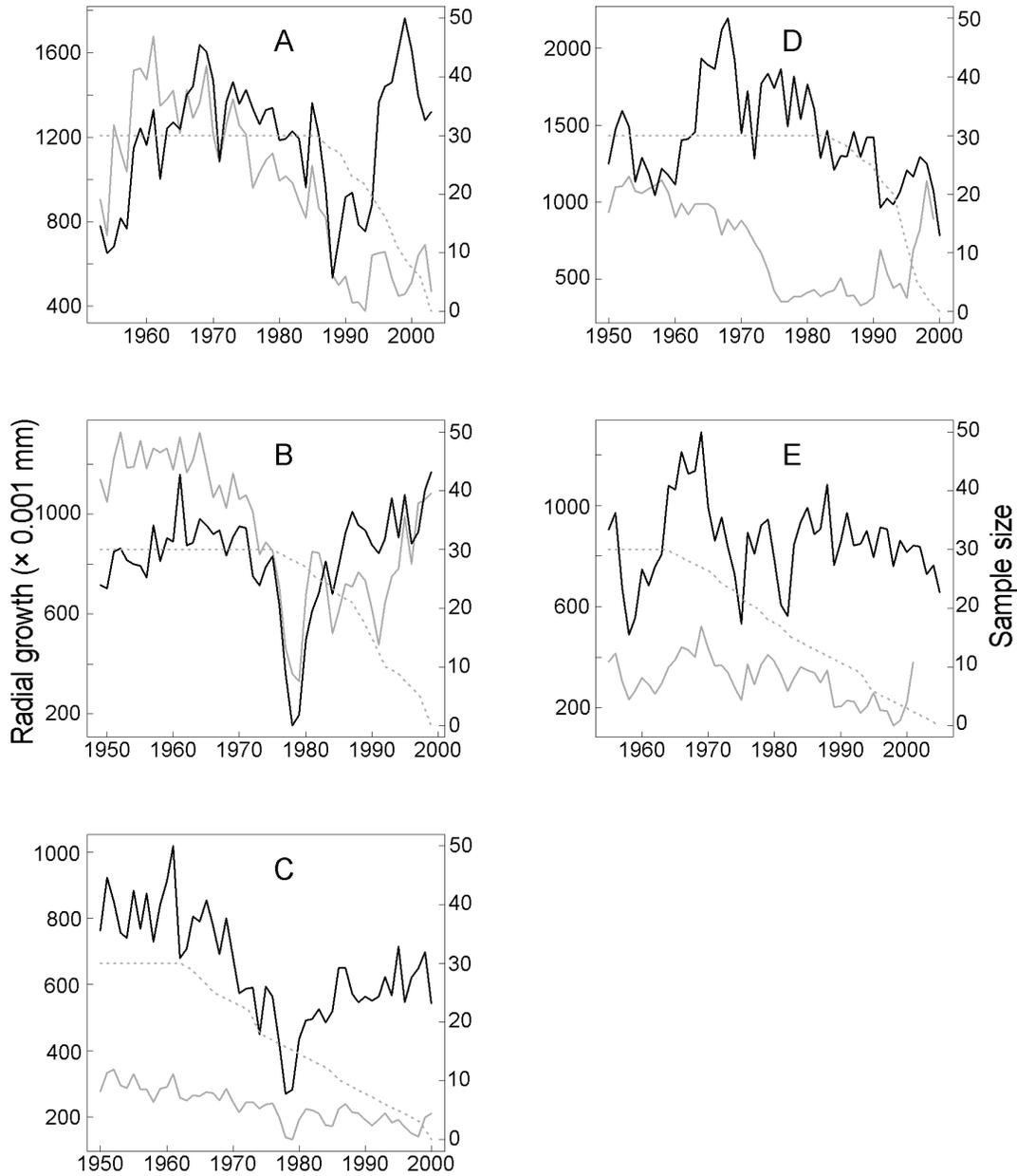

Table 1. Growth series identification, location, growth pattern, tree size (diameter at breast height, DBH), last entire year of growth, and range of year of death of the data sets used as modelling scenarios.

Data set	Location	Growth pattern ~15–20 years prior	DBH range (cm)	Last entire year of growth (live trees)	Range of year of death	Mean years since death
A	Temiscaming, Quebec, Canada	Diverging	19–42	2003	1994–2003	3.9
B	Lower St. Lawrence, Quebec, Canada	Crossover	12–44	1999	1974–1999	18.3
C	Lower St. Lawrence, Quebec, Canada	Diverging	13–37	2000	1963–2000	10.7
D	Abitibi, Quebec, Canada	Merging	na	2000	1979–1999	12.4
E	Pallas Yllas Tunturi, Finland	Merging	na	2005	1938–2001	22.9

bles, which included the slopes calculated over 3, 5, 10, 25, and 35 years of recent growth (SLP_3, SLP_5, SLP_10, SLP_25, and SLP_35, respectively). Considering that the lo-gistic regression analyses used growth level and trend varia-bles, we did not standardize radial growth series nor did we apply any detrending methods. Each of these data treatments

would have eliminated, at least partially, the desired information.

Statistical analysis

We used logistic regression to estimate mortality probabilities as a function of recent radial growth variables. This method is adequate for estimating a binary dependent variable (e.g., live or dead status) from a vector of discrete or continuous variables (Hosmer and Lemeshow 2000). The logistic model is

$$P(Y = 1) = \frac{e^{\alpha G_L + \beta G_T}}{1 + e^{\alpha G_L + \beta G_T}}$$

where $P(Y = 1)$ is the probability that an individual tree dies given the combination of independent variables (where G_L is growth level and G_T is growth trend) and their coefficients α and β (Hosmer and Lemeshow 2000; Quinn and Keough 2002).

Logistic regression analysis was performed in two steps for each data set. The purpose of the first step was to determine which of the many alternative univariate models, each containing one G_L variable, would provide the best fit. Parameter estimation was performed using maximum log-likelihood, and the model yielding the highest likelihood ratio was selected for further analysis. In the second step, G_T variables were added individually to the best univariate model, and a likelihood ratio test was performed to determine whether the bivariate model yielded a statistically significant ($p < 0.05$) improvement to the maximized likelihood over the nested univariate model (Venables and Ripley 2002). In any case, only one G_T and one G_L variable was allowed to remain in the final model. All regression and validation analyses were performed using the statistical computing environment and language R (R Development Core Team 2005).

Once a logistic model has been developed and parameters estimated using measures of goodness of fit and statistical significance as decision criteria, care should be taken to further evaluate the model by assessing prediction error (Fielding and Bell 1997). Solely reporting positive and negative predictive powers of a model can yield a misleading model evaluation, because such measures are sensitive to the prevalence (i.e., the relative frequency of the positive outcome) in the data as well as to the threshold used to determine the binary outcome (Manel et al. 1999). More rigorous measures such as the "receiver operating characteristics" plot, developed from signal detection theory (Mason 1982), can be applied in ecological studies (Guisan and Zimmermann 2000). Hence, we used some of these measures from Harrell's (2001) contributed S-PLUS libraries (Design and Hmisc) (see below) for model assessment.

When model performance is internally evaluated (i.e., using the training (parameterization) data), estimates of the model's predictive accuracy are usually optimistic (Copas 1983; Efron 1983; Van Houwelingen and Le Cessie 1990). This optimism can be accounted for using a bootstrap procedure. First, validation indices are computed for the original data set. Then, at each iteration, these indices are recomputed on the bootstrap sample. Each of the bootstrap indices minus the original indices provides an estimate of the (over-

fitting-induced) optimism of the original fitting process. The optimism indices are averaged over the number of iterations and subtracted from the original indices, yielding a bias-corrected estimate of the predictive accuracy. The bootstrap procedure is a very efficient means of validation, because no data are withheld from the model-fitting process (Steyerberg et al. 2001).

However, external validation is the most stringent means for testing the general applicability of a model to a new data set, and data from a different geographic area may be used to do so (Harrell 2001). To achieve this, the models are first developed and validated with the above-mentioned internal methods, and the resulting best model is then "frozen" and applied to an external data set not related to the training data.

We validated the final "best" model internally using the bootstrap validation procedure validate.lrm available in the R Design package (2.0-12; Harrell 2005), using Somer's D_{XY} index as criteria of model discriminative ability (Somers 1962). The latter is closely related to the area under the curve of the receiver operating characteristics plot (Engelmann et al. 2003) and is, as such, independent of prevalence and classification threshold (Swets 1988). The area under the curve (and, thus, Somer's D_{XY}) is a good indicator in ecological research of responses that are naturally dichotomous, such as the occurrence or nonoccurrence of events such as death (Murtaugh 1996). The D_{XY} values range from –1 (perfect status misclassification, where the model classifies all live trees as dead and vice versa) to 1 (perfect status classification). Values >0.6 indicate a useful modelling application according to Manel et al. (1999). However, we added Nagelkerke's R_N^2 as a measure of discriminative strength to have more information for model comparisons (Harrell 2001).

This internal validation was performed to obtain an unbiased (from overfitting) measure of the discriminative ability of the developed models. However, bootstrapping does not eliminate the underlying growth dynamics in each data set. Hence, external validation on an independent data set was performed to more rigorously test model discriminative abilities. Although we used data set from other tree species, the purpose was only to validate our models on other independent data sets. This external validation should be regarded only as a stringent form of model testing (Harrell 2001) and not an attempt to test whether the model is universally applicable across species. We used the val.prob procedure from the R Desgin package (2.0-12; Harrell 2005) to perform this validation. Predicted and observed probabilities were used to compute several discrimination measures (e.g., D_{XY}) where the selected model of one data set is applied with its original calibration to another.

Results

Truncating synchronizes the distribution of the last year of growth of live trees with the distribution of the year death of dead trees allowing growth comparisons of live and dead individuals to be based on variables computed over the same time period. Depending on the data set, truncating may increase (data set B) or decrease (data set E) values of the pre-

Table 2. Most significant variables, selected variables, and associated internal predictive and discriminative measures (R_N^2 and D_{XY}) of the five data sets used in the modelling procedure based on truncated and untruncated live tree growth series.

Data set	Untruncated			Truncated		
	Final model	R_N^2	D_{XY}	Final model	R_N^2	D_{XY}
A	Med_10 (–0.001 917) + SLP_25 (–0.050 886)	**0.572**	**0.806**	Med_3 (–0.002 641) + SLP_3 (–0.004 504)	0.536	0.761
B	Med_3 (0.004 22) + SLP_25 (–0.141 74)	**0.645**	**0.820**	Med_5 (0.000 256 1) + SLP_35 (–0.072 344 8)	0.296	0.517
C	Med_3 (–0.019 23)	**0.770**	**0.912**	Med_10 (–0.015 96) + SLP_35 (0.162 83)	0.740	0.883
D	Med_5 (–0.003 828) + SLP_35 (0.028 764)	0.513	0.719	Med_3 (–0.003 527) + SLP_3 (–0.012 734)	**0.648**	**0.854**
E	Med_10 (–0.085 07) + SLP_25 (0.713 04)	0.670	0.850	Med_10 (–0.089 57)	**0.722**	**0.897**

Note: Parameter estimates are given in parentheses. Boldface values are the highest score of D_{XY} and associated R_N^2 values between the final truncated and untruncated model.

Table 3. External validation: predictive and discriminative measures (R_N^2 and D_{XY}) of untruncated (UT) and truncated (T) models developed on different training data sets when applied to the various test data sets.

Test data set		Training data set									
		A (merging)		B (cross-over)		C (merging)		D (diverging)		E (diverging)	
		UT	T	UT	T	UT	T	UT	T	UT	T
A (merging)	D_{XY}	—	—	0.376	0.457	**0.731**	0.334	0.540	**0.708**	0.433	**0.631**
	R_N^2	—	—	0.144	0.166	**0.372**	0.179	0.274	**0.518**	0.203	**0.311**
B (cross-over)	D_{XY}	**0.706**	0.260	—	—	0.406	0.099	0.158	0.271	0.163	0.209
	R_N^2	**0.430**	0.105	—	—	0.178	0.001	0.039	0.097	0.019	0.069
C (merging)	D_{XY}	**0.880**	0.830	0.184	0.163	—	—	**0.900**	0.553	0.729	**0.862**
	R_N^2	**0.709**	0.651	0.028	0.002	—	—	**0.754**	0.320	0.493	**0.695**
D (diverging)	D_{XY}	0.162	**0.821**	0.383	0.380	0.252	0.471	—	—	0.603	**0.683**
	R_N^2	0.039	**0.642**	0.177	0.144	0.091	0.202	—	—	0.294	**0.452**
E (diverging)	D_{XY}	0.568	**0.823**	0.559	0.116	0.741	**0.891**	0.813	0.691	—	—
	R_N^2	0.294	**0.526**	0.330	0.025	0.419	**0.739**	0.582	0.393	—	—

Note: Boldface values indicate the best (truncated vs. untruncated) model's discriminative ability (D_{XY}) with associated predictive power (R_N^2) when $D_{XY} > 0.6$.

dictor variables, but often, no consistent pattern within a data set emerges (data not shown).

For each of the five datas ets, the predictor selection of the final truncated model was different from that of the final untruncated model, but there was no obvious pattern in predictor selection (Table 2). There was also no clear pattern in variable selection; truncating did not yield models with either consistently longer or shorter median or trend intervals than untruncated models. In two models (data set C, untruncated, and data set E, truncated) only the G_L variable was retained in the final model. In all other cases, the regression analyses yielded bivariate models. Similarly, parameter estimates showed no obvious pattern; truncating neither consistently increased nor decreased estimates (Table 2).

Internal validation also showed variable patterns in predictive (R_N^2) and discriminative (D_{XY}) ability depending on the situation. For the two data sets with diverging growth chronologies (data sets A and C) and the data set where chronologies crossed over (data set B), truncating lowered R_N^2 and D_{XY} values (Table 2). However, for the two other data sets with merging growth trends, truncated models were of better predictive and discriminative strength (Table 2).

External validation showed that truncating increased the predictive–discriminative strength of the models when they were applied to data sets with different growth patterns (Table 3). Truncated models based on data sets A and C (both merging) performed better on data set D and E (both

diverging) than the untruncated models. However, the cross-over growth trend from data set B did not yield a single good model, i.e., all models had D_{XY} values <0.6 whether or not they were applied to similar (diverging) or dissimilar (merging) data. Untruncated models had higher discriminative ability when run on data with similar growth patterns. Good predictive-discriminative ability ($D_{XY} > 0.6$) was observed when the untruncated model from data set A was run on data set B or C or when the untruncated model of data set C was run on data set A (Table 3, training data sets A to C). Also, the data set C model performed poorly on only B even though the higher D_{XY} and R_N^2 values of the untruncated model seems to indicate that untruncated models are more discriminative when used on data of similar growth dynamics (Table 3, training data set C).

External validation of models based on the diverging growth chronologies (data sets D and E) showed somewhat less obvious results (Table 3). Truncated models from data set E performed better when run on dissimilar data sets A and C (Table 3, training data set column E). However, truncated models of data set D performed better when run on dissimilar data set A, but this was not the case when run on data set C, although the latter is also dissimilar. Similarly, the untruncated data set D model exhibited higher predictive–discriminative values when run on data set E, but the truncated model from data set E had a slightly higher, albeit not very impressive ($D_{XY} = 0.683$), predictive–discriminative ability when run on data set D (Table 3). Models based on these

data (data sets D and E) did not have acceptable values (i.e., >0.6) of predictive-discriminative ability when run on the crossover growth pattern whether truncated or not (Table 3, training data set columns D and E).

Discussion

Truncating affected the values of predictor variables computed from live tree growth series (Fig. 1) and, thus, influenced the selection of variables in the models (Table 2). Consequently, truncating also affected parameter estimates (Table 2) and, therefore, was a data treatment with considerable impact on the modelling process of tree mortality probabilities.

The impact of truncating on a model's predictive–discriminative ability depended on the underlying growth patterns of the training and test data. In general, untruncated models had higher discriminative ability when applied to the same populations from which the models had been derived or when the test (or application) population showed a similar growth pattern. In these cases, higher discriminative abilities stem from the fact that growing conditions have improved since dead individuals died, thereby leading to an increase in the growth levels of live trees. However, when such models were tested on data sets with dissimilar growth patterns, the truncated models were of better predictive–discriminative strength.

Logistic regression uses the difference between groups in the values of predictor variables to predict the probabilities of a binary outcome (e.g., live or dead status). The greater the difference in the predictor variable values, the higher the predictive–discriminative capacity of the resulting model. Here lies one of the reasons why growth trends (e.g., diverging vs. merging) determine the impact of truncating on the model performance. In situations of diverging chronologies, truncating tends to lower the difference in the predictor variable values computed for the live and dead tree populations. Thus, the procedure is expected to lower the predictive–discriminative strength of the model, whereas the opposite is true when chronologies show a merging trend.

Although a high predictive–discriminative strength is desirable when one wants to estimate mortality probabilities, the use of untruncated data in situations where external factors (e.g., disturbance or climate) may have had a potentially strong influence on the recent growth of trees gives this strength a somewhat artificial nature. This would not be the case for dead trees if many of them had already died before the onset of the external influence. External factors, such as disturbance or climate, very often influence tree growth (Lorimer 1985; Nowacki and Abrams 1997), and this may lead to biased mortality models. The use of untruncated data will then confound climate or disturbance growth signals with tree vigour-related growth responses and will lead to high data specificity through the fitting of data-specific noise and, thus, to overfitted models (Harrell 2001). Truncating may alleviate this problem as is suggested by their superior external predictive–discriminative ability.

Our results suggest that truncating reduces the influence of external factors (e.g., disturbance or climate) on mortality probability estimations, which translates into more accurate models with greater general applicability. Therefore, such a procedure is expected to bring the model closer to its original intent, that is, to predict mortality probabilities from radial growth as a surrogate measure of tree vigour (Waring and Pitman 1985; Pedersen 1998a, 1998b). Whether our approach succeeds better in doing so than other approaches, such as the explicit modelling of climate signal and intervention occurrences in growth series (Pedersen 1998b), has not been tested in our analyses and is beyond the scope of this paper. However, truncating has the advantage of being a simple method that does not rely on statistical assumptions and, therefore, that is not influenced by parameter estimation uncertainties.

If one uses growth variables to estimate mortality probabilities in trees, the general use of truncating is indicated. Truncating has shown to improve or, at worst, have little impact on model predictions. However, even lower predictive–discriminative strength in truncated models would promote truncating, because these models more realistically reflect the underlying tree vigour relationships between live and dead trees.

The poor performance of the data set C models in discriminating tree status in other data sets and the poor performance of other models in predicting data set C outcomes highlight the need for cautious parameterization of mortality models. As indicated by the growth inversion following disturbance (Fig. 2B), vigorous and dominant trees were most severely affected by the disturbance agent which was, in this case, the spruce budworm (data set C was derived from balsam fir, the preferred host of the budworm). The subsequent decline and death of the overstory will enhance growth of suppressed and less vigorous trees. Because balsam fir responds very well to release, even after extended periods of suppression (Frank 1990), and repetitively attacked overstory trees die in high proportions (Solomon et al. 2003), the disturbance caused a rapid shift in tree vigour from the dominant canopy trees to the lower strata. This prevented the typical negative growth trend to develop in declining trees; because some of the trees died even though they maintained high radial growth rates (Fig. 2B), neither growth trend nor growth level variables from data set C are good predictors of mortality probabilities in other scenarios.

Tree size, where available, was used as a pairing factor to match live and dead trees to determine the year in which individual live series had to be truncated. Although this might be adequate in most situations, other potential confounding factors, such as tree vigour classes or crown position, could be useful in situations where, as in data set C, vigorous, dominant trees show an abrupt decrease in growth before their death. In this instance, their dominant position may be closely linked to their decline (Batzer and Popp 1985).

Truncated models generally performed better than untruncated models when the growth pattern of the testing data was different. More data spanning an even wider gradient of patterns would be necessary to conduct a more thorough analysis and to pinpoint the causes and extent of this behaviour. Resampling procedures producing confidence intervals of the predictive–discriminative measures could facilitate this task through formal hypothesis testing, but such procedures are currently unavailable (F. Harrell, personal communication). Also, simulated data might be helpful to further investigate the statistical behaviour of truncated models

under a variety of scenarios including multiple disturbances. Such investigations could be useful even for applications where tree mortality probabilities are estimated in a longitudinal ("on-time") approach (Bigler and Bugmann 2004*b*).

Overall, truncating has been shown to be an efficient data treatment in the context of mortality probability estimations. Although more studies may be necessary to further test the impact of truncating on model selection and performance and — in studies where probability estimations are the actual aim — on parameter estimations, it has been demonstrated that truncating can reduce the influence of external factors on the modelling process. When mortality algorithms are to be integrated in forest-simulation models of general applicability, truncating will give the predictors (radial growth variables) more biological sense (Hawkes 2000) and could allow more process-oriented (vigour-oriented) model parameterization, which is an important feature of mortality algorithms especially in a world of accelerated environmental change (Keane et al. 2001).

In simulation models (e.g., JABOWA; Botkin et al. 1972) where mortality occurs when growth becomes lower than a given threshold and where model parameters (i.e., the threshold growth level in the case of mortality prediction) are derived from "intelligent" guesses (Hawkes 2000), truncating does not apply, because no empirical parameter estimation is undertaken. However, these models usually assume that mortality algorithm parameters are constant across species, which is an unrealistic assumption (Kobe et al. 1995; Pacala et al. 1996; Wyckoff and Clark 2002). To counteract these shortcomings, a mortality algorithm has to be parameterized by comparing growth of live and dead trees (Wyckoff and Clark 2000), and truncating in these situations will make the parameter estimates more realistic.

However, we do not propose truncating as a means to develop a general mortality model spanning across species and geographic regions. Mortality probability models derived from field data are empirical in nature and remain dependent on parameterization. Still, this does not diminish the benefits of truncating: making a model as robust as possible where robustness is desired.

Acknowledgements

References

Batzer, H.O., and Popp, M.P. 1985. Forest succession following a spruce budworm outbreak in Minnesota. For. Chron. **61**: 75–80.

Bigler, C., and Bugmann, H. 2003. Growth-dependent mortality models based on tree rings. Can. J. For. Res. **33**: 210–221. doi:10.1139/x02-180.

Bigler, C., and Bugmann, H. 2004*a*. Assessing the performance of theoretical and empirical tree mortality models using tree-ring series of Norway spruce. Ecol. Modell. **174**: 225–239. doi:10. 1016/j.ecolmodel.2003.09.025.

Bigler, C., and Bugmann, H. 2004*b*. Predicting the time of tree death using dendrochronological data. Ecol. Appl. **14**: 902–914. doi:10.1890/03-5011.

Botkin, D.B., Janak, J.F., and Wallis, J.R. 1972. Some ecological consequences of a computer model of forest growth. J. Ecol. **60**: 849–872.

Buchman, R.G. 1983. Survival predictions for major Lake States tree species. USDA For. Serv. North-Central For. Range Exp. Stn. Res. Pap. NC-RP-233.

Buchman, R.G., and Lentz, E.I. 1984. More Lake States tree survival predictions. USDA For. Serv. North-Central For. Range Exp. Stn. Res. Note NC-RN-312.

Copas, J.B. 1983. Regression, prediction and shrinkage. J. R. Stat. Soc. B. **45**: 311–354.

Efron, B. 1983. Estimating the error rate of a prediction rule: some improvements on cross-validation. J. Am. Stat. Assoc. **78**: 316–331. doi:10.2307/2288636.

Engelmann, B., Hayden, E., and Tasche, D. 2003. Testing rating accuracy. Risk, **16**: 82–86.

Fielding, A.H., and Bell, J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. **24**: 38–49. doi:10.1017/ S0376892997000088.

Flewelling, J.W., and Monserud, R.A. 2002. Comparing methods for modelling tree mortality. *In* Proceeding of the 2nd Forest Vegetation Simulator Conference, 12–14 February 2002, Fort Collins, Colo. *Edited by* N.L.Crookston and R.N. Havis. USDA Forest Service, Rocky Mountain Research Station, Fort Collins, Colo. pp. 168–177.

Frank, R.M. 1990. Balsam fir. *In* Silvics of North America. Vol. 2 Hardwoods. *Technical coordinators*: R.M. Burns and B.H. Honkala. US Dep. Agric. Agric. Handb. 654. pp. 26–35.

Franklin, J.F., Shugart, H.H., and Harmon, M.E. 1987. Tree death as an ecological process. Bioscience, **37**: 550–556. doi:10.2307/ 1310665.

Gehrig, M. 2004. Methoden zur Vitalitätsbeurteilung von Baümen. Vergleichende Untersuchungen mit visuellen, nadelanalytischen und bioelektrischen Verfahren. Swiss Federal Institute of Technology, Zürich. Dissertation. No. 15341.

Guisan, A., and Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. Ecol. Modell. **135**: 147–186. doi:10. 1016/S0304-3800(00)00354-9.

Harrell, F.E. 2001. Regression modeling strategies with applications to linear models, logistic regression, and survival analysis. Springer, New York.

Harrell, F.E. 2005. The design package, version 2.0-12 edition. Contributed R package. Available from biostat.mc.vanderbilt. edu/s/Design and biostat.mc.vanderbilt.edu/rms.

Hawkes, C. 2000. Woody plant mortality algorithms: description, problems, and progress. Ecol. Modell. **126**: 225–248. doi:10. 1016/S0304-3800(00)00267-2.

Holmes, R.L. 1983. Computer assisted quality control in tree-ring dating and measurement. Tree-Ring Bull. **43**: 69–78.

Hosmer, D.W., and Lemeshow, S. 2000. Applied logistic regression. 2nd ed. John Wiley & Sons, New York.

Keane, R.E., Austin, M., Field, C., Huth, A., Lexer, M.J., Peters, D., Solomon, A., and Wyckoff, P. 2001. Tree mortality in gap models: application to climate change. Clim. Change, **51**: 509–540. doi:10.1023/A:1012539409854.

Kienholz, R. 1930. The wood structure of a "pistol-butted" mountain hemlock. Am. J. Bot. **17**: 739–764. doi:10.2307/2435684.

Kobe, R.K., Pacala, S.W., Silander, J.A., Jr., and Canham, C.D.

1995. Juvenile tree survivorship as a component of shade tolerance. Ecol. Appl. **5**: 517–532. doi:10.2307/1942040.

Lorimer, C.G. 1985. Methodological considerations in the analysis of forest disturbance history. Can. J. For. Res. **15**: 200–213. doi:10.1139/x85-038.

Manel, S., Dias, J.-M., and Ormerod, S.J. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. Ecol. Modell. **120**: 337–347. doi:10.1016/S0304-3800(99)00113-1.

Manion, P. 1981. Tree disease concepts. Prentice Hall Inc., Englewood Cliffs, N.J.

Mason, I.B. 1982. A model for assessment of weather forecasts. Aust. Meteorol. Mag. **30**: 291–303.

Millers, I., Lachance, D., Burkman, W.D., and Allen, D.C. 1991. North American sugar maple decline project: organization and field methods. USDA For. Serv. Gen. Tech. Rep. NE-154.

Monserud, R.A. 1976. Simulation of forest tree mortality. For. Sci. **22**: 438–444.

Murtaugh, P.A. 1996. The statistical evaluation of ecological indicators. Ecol. Appl. **6**: 132–139. doi:10.2307/2269559.

Nowacki, G.J., and Abrams, M.D. 1997. Radial-growth averaging criteria for reconstructing disturbance histories from presettlement-origin oaks. Ecol. Monogr. **67**: 225–249. doi:10.1890/0012-9615(1997)067[0225:RGACFR]2.0.CO;2.

Ogle, K., Whitham, T.G., and Cobb, N.S. 2000. Tree-ring variation in pinyon predicts likelihood of death following severe drought. Ecology, **81**: 3237–3243. doi:10.1890/0012-9658(2000)081[3237:TRVIPP]2.0.CO;2.

Oliver, C.D., and Larson, B.C. 1996. Forest stand dynamics. Updated edition. John Wiley & Sons, Inc., New York.

Ontario Ministry of Natural Resources (OMNR). 2004. Ontario tree marking guide, version 1.1 edition. Queen's Printer for Ontario, Toronto, Ont.

Ouellet, D., and Zarnovican, R. 1988. Cultural treatment of young yellow birch (*Betula alleghaniensis* Britton) stands: tree classification and stand structure. Can. J. For. Res. **18**: 1581–1586. doi:10.1139/x88-241.

Pacala, S.W., Canham, C.D., Saponara, J., Silander, J.A., Jr., Kobe, R.K., and Ribbens, E. 1996. Forest models defined by field measurements: estimation, error analysis and dynamics. Ecol. Monogr. **66**: 1–43. doi:10.2307/2963479.

Pedersen, B.S. 1998*a*. Modeling tree mortality in response to short- and long-term environmental stresses. Ecol. Modell. **105**: 347–351. doi:10.1016/S0304-3800(97)00162-2.

Pedersen, B.S. 1998*b*. The role of stress in the mortality of midwestern oaks as indicated by growth prior to death. Ecology, **79**: 79–93. doi:10.1890/0012-9658(1998)079[0079:TROSIT]2.0.CO;2.

Peterson, D.W., and Peterson, D.L. 1995. Climatic influences on radial growth of subalpine larch in the north Cascade Mountains. *In* Ecology and management of *Larix* forests: a look ahead. Proceedings of an International Symposium, 5–9 October 1992, Whitefish, Mont. *Compiled by* W.C. Schmidt and K.J. McDonald. USDA For. Serv. Gen. Tech. Rep. GTR-INT-319. pp. 268 – 271.

Quinn, G.P., and Keough, M.J. 2002. Experimental design and data analysis for biologists. Cambridge University Press, Cambridge, UK.

R Development Core Team. 2005. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0. Available from www.R-project.org [accessed 19 April 2005].

Senecal, D., Kneeshaw, D., and Messier, C. 2004. Temporal, spatial, and structural patterns of adult trembling aspen and white spruce mortality in Quebec's boreal forest. Can. J. For. Res. **34**: 396–404. doi:10.1139/x03-263.

Solomon, D.S., Zhang, L., Brann, T.B., and Larrick, D.S. 2003. Mortality patterns following spruce budworm infestation in unprotected spruce–fir forests in Maine. North. J. Appl. For. **20**: 148–153.

Somers, R.H. 1962. A new asymmetric measure of association for ordinal variables. Am. Sociol. Rev. **27**: 799–811. doi:10.2307/2090408.

Steyerberg, E.W., Harrell, F.E., Borsboom, G.J., Eijkemans, M.J.C., Vergouwe, Y., and Habbema, J.D.F. 2001. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J. Clin. Epidemiol. **54**: 774–781. doi:10.1016/S0895-4356(01)00341-9. PMID:11470385.

Swets, J.A. 1988. Measuring the accuracy of diagnostic systems. Science (Washington, D.C.), **240**: 1285–1293. doi:10.1126/science.3287615. PMID:3287615.

Van Houwelingen, J.C., and Le Cessie, S. 1990. Predictive value of statistical models. Stat. Med. **9**: 1303–1325. doi:10.1002/sim.4780091109. PMID:2277880.

Van Mantgem, P.J., Stephenson, N.L., Mutch, L.S., Johnson, V.G., Esperanza, A.M., and Parsons, D.J. 2003. Growth rate predicts mortality of *Abies concolor* in both burned and unburned stands. Can. J. For. Res. **33**: 1029–1038. doi:10.1139/x03-019.

Venables, W.N., and Ripley, B.D. 2002. Modern applied statistics with S-PLUS. 4th ed. Springer, Berlin.

Waring, R.H. 1987. Characteristics of trees predisposed to die. Bioscience, **37**: 569–574. doi:10.2307/1310667.

Waring, R.H., and Pitman, G.B. 1985. Modifying lodgepole pine stands to change susceptibility to mountain pine beetle attack. Ecology, **66**: 889–897. doi:10.2307/1940551.

Wyckoff, P.H., and Clark, J.S. 2000. Predicting tree mortality from diameter growth: a comparison of maximum likelihood and Bayesian approaches. Can. J. For. Res. **30**: 156–167. doi:10.1139/cjfr-30-1-156.

Wyckoff, P.H., and Clark, J.S. 2002. The relationship between growth and mortality for seven co-occurring tree species in the southern Appalachian Mountains. J. Ecol. **90**: 604–615. doi:10.1046/j.1365-2745.2002.00691.x.