Forest Science (2025) 71:697–719
https://doi.org/10.1007/s44391-025-00028-6

ORIGINAL ARTICLE

A Generalized Marginal and Mixed-Effect Models for Predicting Tree-Level Mortality with Unequal Measurement Intervals for Scots Pine in Finland

Jouni Siipilehto[1] · **Harri Mäkinen**[1]

Received: 8 October 2024 / Accepted: 3 April 2025 / Published online: 23 June 2025

Abstract

Logistic regression has been used to model individual tree mortality. However, unequal measurement intervals limit the use of reasonable link functions. Although marginal and mixed-effects models have been used, a comparison of these models for predicting mortality in new stands is lacking. We developed models for predicting the mortality of Scots pine (*Pinus sylvestris*) trees in Finland. The modelling data comprised 44 thinning experiments (127,057 tree-level observations), and 50% of the experiments were randomly selected for model evaluation (43 experiments, 112,518 tree-level observations). A complementary log–log model was used to predict mortality probability based on tree and stand characteristics and thinning effects. The measurement period length was added as an offset variable. The marginal (population-average) and mixed models with random effects (site, plot, and year) were fitted and evaluated. The evaluation consisted of fit statistics, comparisons of predicted and observed mortality rates, and simulations using the Motti stand simulator. The mixed models clearly provided better statistical fits than the marginal model. The evaluation with the Motti simulator showed the most accurate prediction in terms of stem number (N) and stand basal area (G) when using the marginal model. All models, except the random year effect, resulted in maximum G values that remained at a reasonable level in the prolonged Motti simulations. The current survival model in the Motti simulator is based on a relative density index derived from the self-thinning line and provides good prediction accuracy. Based on the new *BAL*-based models, we recommend the marginal model as an option to the current model.

Keywords *Pinus sylvestris* · Between-tree competition · Complementary log–log regression · Generalized linear mixed model · Stand dynamics

---

Extended author information available on the last page of the article

Introduction

Tree mortality may differ in terms of scale and frequency. Tree mortality is often divided into three major classes: *i*) regular mortality related to between-tree competition, ageing, and normal incidence of pests, diseases, and weather phenomena, *ii*) anthropogenic, that is, human-induced mortality mainly caused by harvesting, and *iii*) catastrophic irregular mortality caused by storms, major insect and disease outbreaks, drought, flooding, or other extraordinary large-scale events (Vanclay 1994). Between-tree competition is the most common cause of tree mortality in Finnish forests (Yli-Kojola 2005). In general, relatively small trees have a high probability of mortality due to competition, whereas large trees are prone to insect and wind damage (Laarmann et al. 2009; Sims et al. 2014). Hamilton (1986) classified the factors affecting tree survival into four groups: tree size, tree competition status in a stand, tree viability, and stand density. Tree-level mortality has been predicted using variables such as tree growth, crown and stem dimensions, dominance, and competition, as well as stand, site, and environmental variables (Monserud 1976; Monserud and Sterba 1999; Salas-Eljatib and Weiskittel 2020). Species specific models for total (regular and irregular) tree mortality are needed for reliable long-term simulation of a stand development.

The most common method used to model the probability of tree-level mortality or survival is the logistic regression. Logistic regression with logit link function is a valid option as long as the measurement interval is constant (Yang and Huang 2013). However, when the measurement period length (L) varies, using the logit link with an offset variable for the period length as $\ln(L)$ does not follow the theory of period length as an interest rate (exponent) in a logistic regression (Monserud 1976), but instead becomes a multiplier as shown in Eq. 8 in Yang and Huang (2013). In addition to treating L as an offset variable, there have been attempts to estimate the effect of varying period length by using it as a covariate in logistic regression models (Kiernan et al. 2009; Ma et al. 2013). If the complementary log–log (CLL) link function for tree mortality is chosen instead, period length can be correctly included as an offset variable in logarithmic form, as shown by Fortin et al. (2008).

Typically, published mortality or survival models are marginal models over the entire dataset, without any random variables. The benefit of a marginal model is the unbiased residuals in the modelling data, and the model represents the average behaviour of the entire dataset. Adding random effects improves model fit statistics, such as log-likelihood and Akaike information criterion (AIC). Various random effects can be used depending on the hierarchical nature of the data: experimental area (site), plot and year (Boeck et al. 2014). The random site effect accounts for differences in locations and site types, the random plot effect accounts for unobserved sustained plot effects, such as soil quality or genetic origin within sites, and the random year effect accounts for differences in weather conditions (temperature, precipitation, wind, and global warming) that may have affected catastrophic (irregular) tree mortality (see Fortin et al. 2008). In addition to a random intercept, a random effect can also be associated with a certain covariate. For example, Yang and Huang (2013) applied a random plot effect to the intercept and slope associated with a diameter increment. Despite a better statistical fit, a mixed model may not provide better predictions when simulating stand development, because the fixed part of the mixed-effects model includes some bias over the entire data, especially in unbalanced data. However, a mixed model ensures unbiased tests for the estimated parameters (Goldstein 1995; Jutras et al. 2003). Even when the marginal (population-average) models were used to explore the model structure, the marginal models were not evaluated together with the final mixed-effects models (e.g., Jutras et al. 2003; Siipilehto et al. 2021). When the marginal and mixed-effects models were compared, the random effects were also estimated (Ma et al. 2013). In most practical applications, the mixed-effect models have to be applied as a fixed-effects model for a new stand, as there are no tree-level observations to predict the random effects (Myllymäki et al. 2024). Finally, to have a fair comparison between the estimated marginal and mixed-effects models and to avoid overoptimistic fit characteristics, the models should be tested with an independent test dataset (Boeck et al. 2014).

The self-thinning line shows the maximum density of a stand (*Nmax*) with respect to the mean diameter, e.g. quadratic mean diameter or basal area median diameter (Reineke 1933; Hynynen 1993). The current survival models for different tree species in the Finnish Motti stand simulator (Hynynen et al. 2014) are based on models for self-thinning lines and their calculated relative density factor, *RDF* for stand level competition and *RDFL* for tree-level competition calculated from trees larger than a target tree (Hynynen 1993; Hynynen et al. 2002). The unitless *RDF* is set to 1 when the stand reaches maximum density according to the self-thinning model. For the smallest tree the *RDFL* is equal to the total stand *RDF*, while for the largest tree the *RDFL* is 0. Instead of competition described as relative density, the basal area in larger trees (*BAL*) (Wykoff 1990) has often been used when modelling mortality or survival (Monserud and Sterba 1999; Eid and Tuhus 2001; Fridman and Ståhl 2001). Recently developed models for Norway spruce (*Picea abies* (L.) Karst.) based on *BAL* (Siipilehto et al. 2021) as alternative models for the Motti simulator have shown reliable results. Here we tested whether the widely used *BAL* was better than *RDFL* at predicting Scots pine mortality.

The aim of this study was to improve the model for the probability of mortality of individual Scots pine trees using the generalised marginal and alternative mixed-effects models with unequal measurement period lengths. The models were based on tree size, tree- and stand-level competition, developmental stage, and stand management. We compared (*i*) the marginal model with alternative mixed-effects models, such as the random intercept model for (*ii*) sites, (*iii*) plots nested within the sites, and (*iv*) years nested within the sites. The evaluation consisted of fit statistics (log-likelihood, AIC) and predicted mortality in the test dataset. In addition, we simulated the development of stand density in the unthinned plots of the test dataset to compare the alternative mortality models developed in this study with each other and with the current unpublished model in the Motti stand simulator. The best model developed in this study is intended to be a *BAL*-based option for the current model in Motti, which is based on the estimated self-thinning line (Hynynen 1993; Hynynen et al. 2002). Potential users of the Motti simulator comes from research and practical forestry (Salminen et al. 2005). Nowadays Motti is the most widely used simulator in Finland, as it is used for forest management planning for private forest estates throughout Finland. Motti is also used for scenario analysis of different management intensities (e.g., Hynynen et al. 2014).

Material and Methods

Modelling and Test Datasets

The modelling and test datasets included thinned and unthinned plots from two growth and yield experiment series, ARPVANHA and HARKAS in Finland. Some HARKAS experiments were established in the 1960 s, but most were established in the 1970 s (Mäkinen and Isomäki 2004). Some younger HARKAS experiments were established in the 1990 s. The older ARPVANHA experiments were established in the 1920 s and the 1930 s (Ilvessalo 1932). The stands were pure or almost pure Scots pine stands growing on mineral soil. Forest site types ranged from the grove-like *Oxalis-Myrtillus* type (OMT) to the dry *Calluna* type (CT) according to Cajander's (1926) classification. Forests were mainly naturally regenerated or sown with seeds of local origin, and few stands were planted. The actual planting density was not known, but the recommended density was 2000–2500 trees per hectare. The treatment schedules included unthinned control plots and plots thinned from below, with intensities ranging from low-intensity thinning (6%–20% removal of stand basal area) to 54% removal. Some results based on older experiments have also been published (Nyyssönen 1950; Vuokila 1962; Mielikäinen 1979). Experiments were randomly divided for modelling and test datasets and their locations are shown in Fig. 1.

The mean measurement period was nine years, (range 3–22 years) (Table 1). The mean stand characteristics varied only slightly between the modelling and test datasets. Within the datasets, the variation in stand density was wide. For example, the stand basal area (*G*) varied between 3.8 m^2 ha^{-1} and 46.8 m^2 ha^{-1} (Table 1). A total of 5,939 trees died and 106,579 trees survived in the test dataset, whereas 6,850 trees died and 120,207 survived in the modelling data, resulting in mortality rates of 5.3% and 5.4%, respectively. The number of sites was 43 and 44 in the modelling and test datasets, respectively. The corresponding numbers of stand-level observations, e.g. *G*, stem number (*N*), basal area-weighted mean diameter (*DG*), and quadratic mean diameter (*DQ*) were 1,084 and 822, respectively.

Complementary Log–Log Model for Tree Mortality

When modelling individual tree mortality, the uneven lengths of the measurement periods had to be taken into account. The probability of tree mortality was expressed as an exponential function in which the period length was included in the exponent as an interest rate (Fortin et al. 2008):

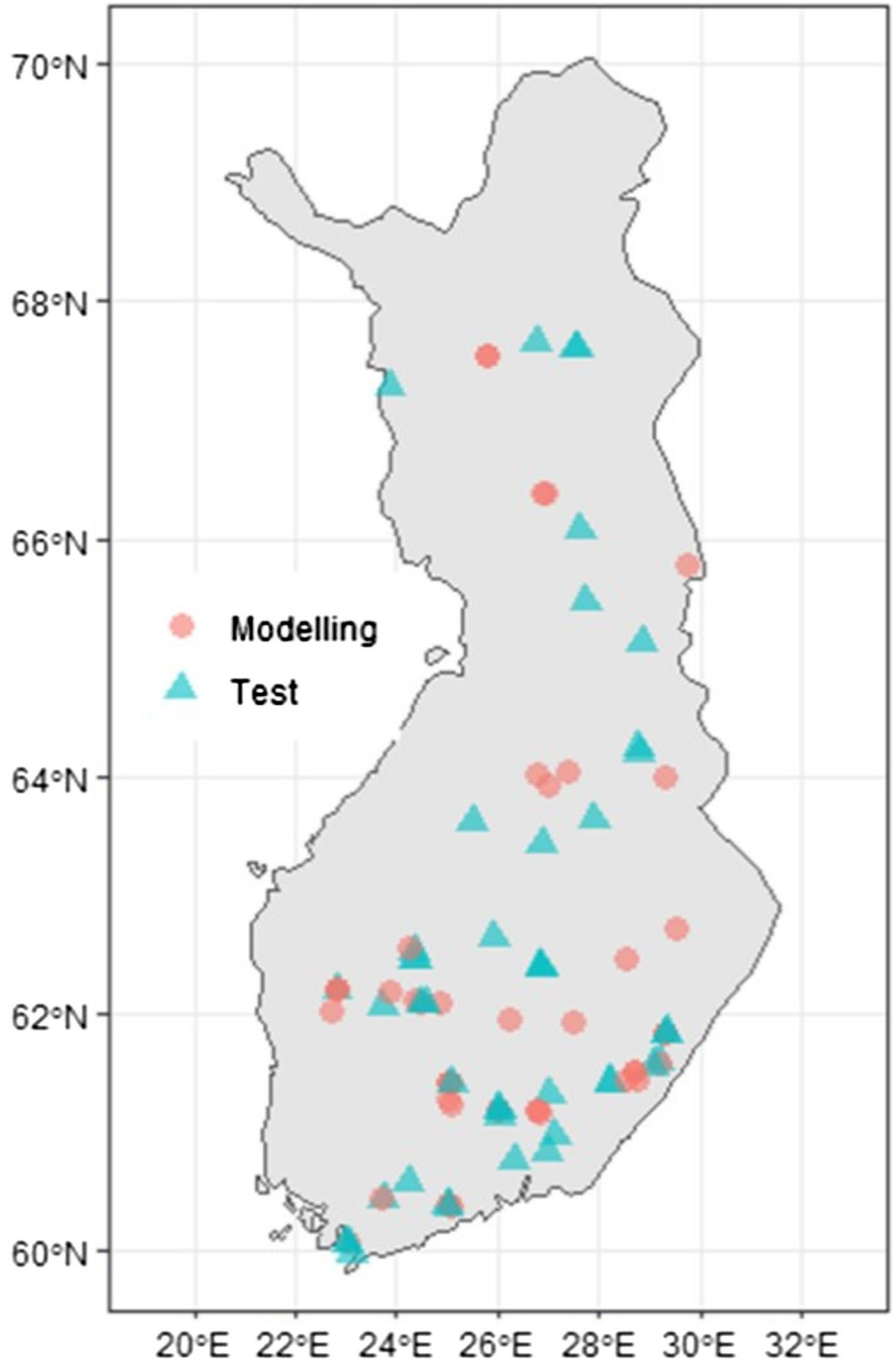

Fig. 1 Locations of Scots pine growth and yield experiments in Finland used to estimate mortality model parameters (modelling) and assess model performance (test)

Table 1 Site, stand, and tree characteristics in the modelling (randomly selected 50% of the experiments) and test datasets for Scots pine mortality in Finland. All the tree and stand characteristics represented values at the beginning of the growth period

	Modelling data			Test data		
Variable	Mean	Min	Max	Mean	Min	Max
Site level						
Plot area, m^2	1233	314	2500	1179	600	2500
GDD5	1153	692	1405	1126	641	1399
Longitude, °	26.2	22.7	29.8	26.2	22.8	29.3
Latitude, °	62.4	60.1	67.5	62.5	60.0	67.6
Altitude, m	130.7	24.0	299.0	130.9	7.0	375.0
Period length, yrs	9.1	3	22	8.7	3	19
Plot level						
Age, yrs	54.0	20	124	51.1	18	103
G, $m^2\ ha^{-1}$	20.1	4.1	46.3	20.9	3.8	46.8
N, ha^{-1}	1012	145	5159	1110	83	4520
DG, cm	18.1	8.0	41.5	18.0	7.5	44.6
DQ, cm	17.1	6.9	40.2	17.0	7.0	43.3
Thinned, proportion	0.67	0	1	0.73	0	1
Tree level						
dbh, cm	14.5	1.0	56.3	15.2	1.0	41.6
BAL, $m^2\ ha^{-1}$	14.1	0.1	57.8	13.9	0.1	62.1
Mortality, proportion	0.054	0	1	0.053	0	1

G is basal area, *N* is stem number, *DG* is basal area-weighted mean diameter, *DQ* is quadratic mean diameter, *BAL* is basal area in larger trees, *dbh* is stem diameter at breast height (1.3 m), *GDD5* is the average growing degree days using the 5 °C threshold for a 30-year period from 1971 to 2000 from the ClimateDT service

$$p = 1 - \exp(-\exp(\mathbf{X}'b))^{L} \tag{1}$$

where p is the probability of tree mortality and $\boldsymbol{X}'b$ is a linear model: a transposed vector of explanatory variables $\boldsymbol{X}$ and a vector of estimated parameters b. The measurement period length (L) varied substantially, from 3 to 22 years, and therefore, the mortality model was fitted using the complementary log–log (CLL) link in Eq. (1) (e.g., Fortin et al. 2008; Myllymäki et al. 2024):

$$CLL(p) = \ln(-\ln(1-p)) = X'b + \ln(L) \tag{2}$$

where the variation in period length (L) was an offset variable (i.e., the coefficient was one) in the linear Eq. (2). Note that the option for modelling mortality ("failures") with CLL would be modelling survival ("successes") using the log–log link function, but then we need to use the negative form of the period length -ln(L) as an offset variable. Yang and Huang (2013) used the logit link function, but then the period length ln(L) was not in accordance with the theory of period length affecting as interest rate.

Random Effect in the Model

The random effects model with binary response (e.g., Stiratelli et al. 1984) used for modelling tree mortality was first presented by Jutras et al. (2003). This approach is also called multilevel model with binary response. The idea is to describe the stand-level dynamics instead of a population-average over the entire dataset. The advantage of the random effect model is that it correctly addresses the hierarchical structure of the data, and it ensures unbiased tests for all model parameters. In the HARKAS and ARPVANHA datasets, tree and stand level characteristics are based on treatment plots (i.e., randomised thinning intensity or unthinned control) within each experiment (or site for now on). The observed response is represented by the binary variable $y_{ijk,}$ the status of the i^{th} tree of the j^{th} plot in the k^{th} site at the end of each measurement period (0 = alive, 1 = dead). The distribution of the response is $y_{ijk} \sim \text{Bin}(p_{ijk},1)$. The model with a random effect on the intercept term can be written in general form as follows:

$$p_{ijk} = f\left(\boldsymbol{X'}_{ijk}b + u_{jk}\right), \tag{3}$$

where p_{ijk} is the expectation of the response for i^{th} tree on plot j at site k, vector $\mathbf{X'}_{\text{ijk}}$ is the transposed vector of independent fixed variables, b is the vector of the estimated parameters, and u_{jk} represents the random variables for site k or nested plot j within site k (i.e., plot(site)) or nested calendar year of the measurement j within site k (i.e., year(site)) with $u \sim \text{N}(0,{s_u}^2)$. The tree-level variance was assumed to be 1 and tested for over- or under-dispersion (Goldstein 1995).

The generalised mixed-effect model was fitted using the GLIMMIX procedure with quasi-Newton optimization in SAS (version 9.4, SAS Institute Inc. 2017). Using the Laplace method, the starting values for the parameter estimates for mixed models were selected from the estimated marginal model. In contrast to the nonlinear NLMIXED procedure with only one possible random effect (Siipilehto et al. 2021), GLIMMIX enables a flexible combination of different random effects.

Model Formulation

Although tree size is not a direct cause of death, the main predictor variables used in the models were the stem diameter and competition status of a tree. In addition, we considered the ecologically and biologically important effects of developmental status, stand density, stand structure, and site factors (Yang et al. 2003). Basal area in larger trees (*BAL*) is a widely used competition measure. *BAL* can also be used as an index value by dividing it by the total stand basal area G, as $BALindex = BAL/G$ (Vanclay 1994; Jutras et al. 2003). In the tree-level survival model for Norway spruce, Siipilehto et al. (2021) found that *BAL* divided by square root of breast height diameter ($\sqrt{dbh}$) was a superior independent variable compared with other *BAL*-based variables. Thus, we tested BAL, BAL/dbh, and $BAL/\sqrt{dbh}$ as well as dbh, $1/dbh$, dbh^2, $dbh^{1.5}$, and $dbh^{1.2}$, as tree-level predictors. A variable dbh^2 is often needed to predict the U-shaped mortality trend caused by increasing mortality among the largest trees of Norway spruce, and it may or may not be needed to predict pine mortality (Monserud and Sterba 1999; Yao et al. 2001). We also tried lower powers of 1.5 or 1.2 to avoid a too strong U-shape.

The development stage can be described by age or other mean stand characteristics, such as the dominant height (*Hdom*), basal-area-weighted mean diameter (*DG*), basal-area-weighted mean height (*HG*), or quadratic mean diameter (*DQ*). The variable *DQ* is commonly used when modelling a stand structure (Newton et al. 2005; Coble and Lee 2008), especially when modelling self-thinning (Monserud et al. 2005; Weiskittel et al. 2009). It can be calculated as $DQ = \sqrt{(G/qN)}$, where G is the stand basal area, and q is the conversion factor $q = \pi/200^2$. The mean characteristics are highly correlated with each other; therefore, only one or a few of them are usually included.

Most of the plots in the HARKAS and ARPVANHA experiments were thinned. Thinning from below most likely decreases the observed mortality probability of a tree because the most suppressed trees are removed. We included the effect of thinning in the models using dummy variables describing the time since the last thinning, ranging from 0–5 years, 6–10 years, and 11–15 years after the treatment. A dummy variable (also known as an indicator variable) takes a binary value (0 or 1) to indicate whether the plot was thinned within the specified time (1) or not (0).

The final model is as follows:

$$X\prime b = f\left(dbh, dbh^{1.2}, \frac{BAL}{\sqrt{dbh}}, \ln(DQ), \ln(G), th\_0\_5, th\_6\_10\right) \tag{4}$$

where *dbh* is the stem diameter at breast height (cm), *BAL* is the basal area in larger trees (m^2 ha^{-1}), *DQ* is the quadratic mean diameter (cm), *G* is the basal area (m^2 ha^{-1}), and *th_0_5* and *th_6_10* are dummy variables for time since the last thinning for 0–5 years and 6–10 years. All the tree and stand variables represented values at the beginning of the growth period.

For comparison, we also tested the current (unpublished) survival model in Motti. The model is expressed as follows:

$$X\prime b = 8.8165 - 4.4671RDFL - 55.5793RDFL/dsh \tag{5}$$

where *RDFL* is the competition index (relative density factor) calculated from trees larger than the target tree, and *dsh* is the stem diameter estimated to stump height (height of the uppermost root neck) using the equation by Laasasenaho (1975) as $dsh = 2.0 + 1.25\ dbh$. The variable *RDFL* is based on the self-thinning line estimated by Hynynen (1993) and is described in detail by Hynynen et al. (2002). The explanatory variables of the current survival model represent values after the growth period. The current survival model was estimated in 2010 using NLMIXED in SAS based on the data from the HARKAS experiments measured by then, including the random site effect in the nonlinear logistic equation.

Evaluation using the Motti Simulator

The test dataset consisted of ten unthinned plots, which were used as an evaluation dataset for the simulations with the Motti stand simulator. Motti is a comprehensive analysis tool and decision support system for assessing the impacts of forest management alternatives on stand dynamics (Salminen et al. 2005; Hynynen et al. 2014). The unthinned plots were monitored for a period of between of 16 and 31 years. The stand-level characteristics of the first measurement of each plot was used as an initial state, with the length of each simulation period aligned with that of the length of the measurement periods. In Motti, the stand level information was converted into tree lists using parameter recovery of the Weibull function for sampling tree dbh (Siipilehto and Mehtätalo 2013) and Näslund's height model for tree height (Siipilehto and Kangas 2015). The predicted tree list includes the knowledge of the number of trees that the tree record represents in the respective stand. The predicted probability of tree mortality is used to continuously and smoothly reduce the number of trees/ha represented by each sample tree (see Monserud and Sterba 1999; Ahtikoski et al. 2018). The predictions for the number of stems ha^{-1} (N) and stand basal area $\mathrm{m}^2\ \mathrm{ha}^{-1}$ (G) using the *BAL*-based marginal and mixed-effects models developed in this study were compared with the measurements as well as with the predictions of the current survival model used in the Motti simulator. Ultimately, the simulations were extended until the maximum G was achieved to ascertain whether the tree-level model could effectively limit the stand basal area to a reasonable level.

Results

The Estimated Mortality Models

We conducted a comparative analysis of the marginal model and the mixed-effects models with different random effects for Scots pine mortality. The estimated parameters of the fitted marginal and mixed-effects models were found to be statistically highly significant ($P < 0.0001$), with the exception of $\ln(G)$ in the random site model ($P = 0.02$) and *th_0_5* in the random plot model ($P = 0.12$). There was considerable variation in the estimated parameters between the alternative models. The impact of *BAL* in the random plot model was considerably less pronounced than in the other models, while it was more driven by *dbh* and G (Table 2). The most notable discrepancy was the negative parameter for $\ln(G)$ in the random year model. However, the mixed model with the random year effect demonstrated superior fit statistics, as evidenced by the − 2 loglikelihood and AIC values, (Table 2). Nevertheless, we proceeded with further evaluation of this model, despite the somewhat illogical effect of $\ln(G)$. There was a significant improvement in the fit statistics when moving from the marginal model to the mixed models with random site (2278 units), and then from the random site to the plot effect model (742 units). Finally, there was a further improvement when moving from the plot effect to the random year effect model (still 538 units) (Table 2). The discrepancies between the mortality models were analysed using an extreme stand with the following characteristics: $N = 500$

Table 2 The estimated parameters of the marginal and mixed-effects complementary log–log models for Scots pine mortality, along with the fit statistics (− 2 log likelihood and AIC). The estimated random effect is shown in the model name: random site, nested random effects as plot(site), and year(site)

Model	Marginal		Random Site		Random plot(Site)		Random year(Site)	
Variable	Estimate	Std	Estimate	Std	Estimate	Std	Estimate	Std
Intercept	− 9.0177	0.2998	− 14.337	0.4416	− 1.4163	0.6376	− 9.8060	0.3640
dbh	− 0.8894	0.0660	− 1.9143	9.9796	− 1.7804	0.0736	− 1.4519	0.0717
$dbh^{1.2}$	0.3525	0.0310	0.8068	0.0354	0.7717	0.0334	0.6135	0.0325
$BAL/\sqrt{dbh}$	0.1832	0.0110	0.0719	0.0184	0.1621	0.0165	0.1822	0.0152
ln(*DQ*)	2.1962	0.1156	3.4899	0.1754	1.0391	0.2081	3.0468	0.1419
ln(*G*)	0.3135	0.0910	1.9786	0.1666	− 0.4201	0.0995	0.2370	0.1017
th_0_5	− 0.4996	0.0396	− 0.0930	0.0605	− 0.5200	0.0522	− 0.5339	0.0447
th_6_10	− 0.3559	0.0601	− 0.1909	0.0748	− 0.4039	0.0733	− 0.3955	0.0601
Variance	1.0835		1.0507	0.1156	1.1787	0.1692	0.4604	0.1043
− 2 loglike	35,670		32,649		32,110		33,391	
AIC	35,686		32,666		32,128		33,409	

* *dbh*= stem diameter at breast height (1.3 m), *BAL*= basal area in larger trees (m^2 ha^{-1}), *DQ*= quadratic mean diameter (cm), *G*= basal area (m^2 ha^{-1}), *th_0_5* and *th_6_10* are dummy variables that take the value 1 if the time since the last thinning was 0–5 years or 6–10 years respectively, and the value 0 otherwise

ha^{-1}, *G*= 45 m^2 ha^{-1}, and *DQ*= 33.8 cm. The curves are presented up to a *dbh* of 60 cm, which is beyond the dataset. Following this examination, it became evident that the mortality profiles were rather different (Fig. 2). The profiles for the random plot and the random year model were found to be disparate (Fig. 2). It is notable that the mortality rate for the largest trees increased from the marginal model, where it was almost non-existent (0.4%) despite the significant parameter for $dbh^{1.2}$, to the random site model (14%), the random year model (24%), and finally the random plot model (60%).

Model Evaluation

Observed and Predicted Mortality Rates in the Test Dataset

The predicted mortality rates are examined over the entire test data and for classes of tree-, plot- and site-level characteristics in test dataset. Our analysis revealed a minor underestimation in the mean predicted mortality for the random site and random plot models (4.2–4.0%) compared to the observed mortality of 5.3% in the test dataset. The random year model produced a slight underestimation of 4.8%, while the marginal model provided an almost accurate mortality prediction of 5.2% for the entire test dataset.

In the smallest diameter class (2.5 cm), the predictions of the alternative models exhibited considerable variation in test dataset (Fig. 3A). The mixed model with the random plot effect predicted the lowest mortality rate (27%) for the test dataset,

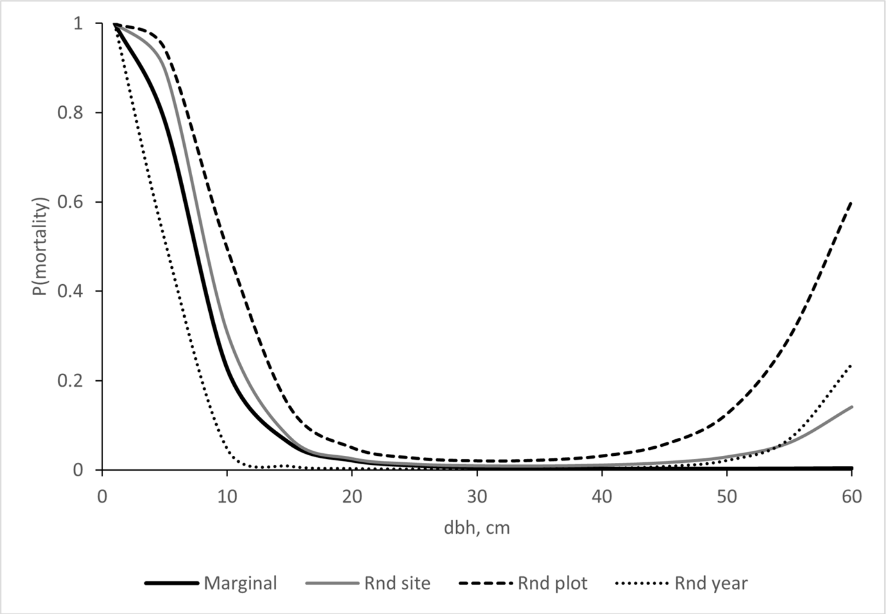

Fig. 2 The U-shaped mortality profiles predicted by the different models with respect to tree *dbh* when the stem number $N = 500$ ha^{-1}, basal area $G = 45$ m^{2}ha^{-1}, and quadratic mean diameter $DQ = 33.8$ cm

while the mixed model with the random year effect exhibited the highest mortality rate (62%). The observed mortality rate of 32% in the test dataset was most closely predicted by the random site effect (32%) and marginal models (31%). In the second-smallest diameter class (7.5 cm), the mixed model with a random year effect slightly overestimated the observed 13.4% mortality rate in the test dataset (16.5%) while the other models slightly underestimated it, least by the marginal model (Fig. 3A). In the *dbh* classes exceeding 17.5 cm, the lowest mortality rates were provided by the random year effect model (Fig. 3A). It underestimated the mortality for the 17.5 and 22.5 cm classes but provided the best predictions for the 27.5 and 32.5 cm classes. It was found that the mortality rate of the largest trees was increasing in test dataset. The random site and random plot models overestimated mortality for the largest trees while the marginal and random year models provided quite accurate predictions (Fig. 3A). The upward trend in predicted mortality rates was attributable to the negative parameter for $dbh^{1.2}$.

Upon examination of mortality in relation to the *BAL* classes, the predicted mortality rates were generally underestimated by the models from the 22.5 *BAL* class to higher *BAL* classes (Fig. 3B). As anticipated, the observed mortality rate generally increased in line with the *BAL* (Fig. 3B). However, the increase was not continuous, with the highest mortality rates observed in the 37.5 m^2 ha^{-1} *BAL* class, and unexpectedly lower in the largest *BAL* class. It was therefore to be expected that there would be inaccuracy in the predicted mortality rates against *BAL*. We opted to utilise a single *BAL*-based variable, and the $BAL/\sqrt{(\text{dbh})}$ proved to be the most suitable option. The marginal model predicted the highest

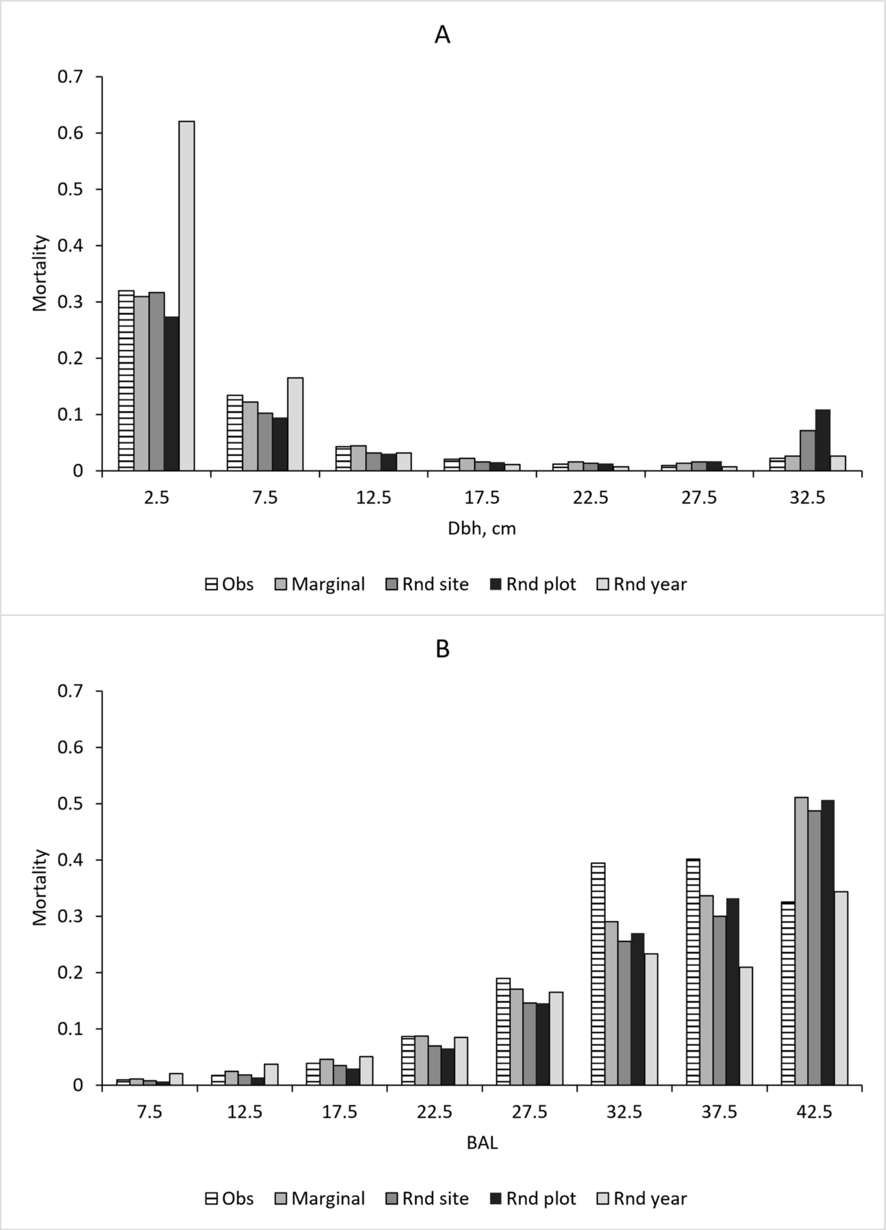

Fig. 3 The observed (obs) and predicted survival rates in the *dbh* and the basal area in larger trees (*BAL*) classes in the test dataset using the marginal and mixed-effects mortality models with random site, plot, and year effects. Note that the y-axis is capped at a 0.7 mortality rate

mortality rates, and generally, they were closest to the observed mortality rates, except for the highest 42.5 m^2 ha^{-1} *BAL* class, in which the random year model provided the best prediction in the test dataset (Fig. 3B).

The marginal model proved to be the most accurate in predicting the mortality rate against quadratic mean diameter (*DQ*) (Fig. 4A). However, the random site model provided slightly more accurate predictions for the 7.5 and 22.5 cm *DQ* classes in the test dataset. The random year model overestimated mortality in the smallest *DQ* class of 7.5 cm, while its predictions for the higher *DQ* classes (17.5–27.5 cm) were clearly underestimates. The random site and the random plot models provided almost equal predictions except for the smallest and highest *DQ* classes.

The test data indicate that the highest basal area did not correspond to the highest mortality rate (Fig. 4B). However, the highest level of predicted mortality was found in the highest *G* class. The random year model provided the best prediction for the highest *G* class. The random year model produced clear overestimates for the two lowest basal area classes, while underestimating mortality for the higher basal area classes of 32.5 and 37.5 m^2 ha^{-1}. In general, the marginal model provided a reasonable prediction for the whole range of *G*.

The test dataset did not indicate a clear increasing trend in mortality with increasing age (Fig. 5A). In the case of the youngest stands, the random year model provided a clear overestimation of mortality (14.5%), which is in stark contrast to the 1% mortality rate observed in the test dataset (Fig. 5A). In this age group, the random plot model provided the most accurate predictions. However, as age increased, the random plot model generally underestimated mortality. The exception was the oldest age class (110), in which the random plot model overestimated mortality the most (Fig. 5A). For the whole range of age the marginal model provided reasonable predictions of mortality and between 50 to 90 years age classes the marginal model provided the most accurate predictions.

Overall, the models provided accurate predictions with respect to years since last thinning in the test dataset (Fig. 5B). However, except for the random year model, the models clearly overestimated mortality for the class of 35 years since last thinning. The random year model provided the most accurate predictions for the unthinned stands and for the classes of 15 years and 35 years since last thinning (Fig. 5B). The marginal model slightly overestimated mortality in the thinned stands. For 5 to 25 years since last thinning, the random site model provided almost correct mortality rates. For the unthinned stands, the random plot model clearly underestimated mortality, while the marginal model and the random year model demonstrated the highest prediction accuracy (Fig. 5B).

In the test dataset, all the models overestimated mortality in northernmost Finland, 600 *GDD5* class (Fig. 6). The test data showed an exceptionally high mortality rate of 10% for the 800 *GDD5* class. However, the predicted mortality was overestimated by the marginal model and accurately predicted by the random site model (Fig. 6). As this higher mortality could be predicted, it was not influenced by exceptional irregular mortality but, was due to experiments with a lower proportion of thinned stand plots. The random site model provided the most accurate

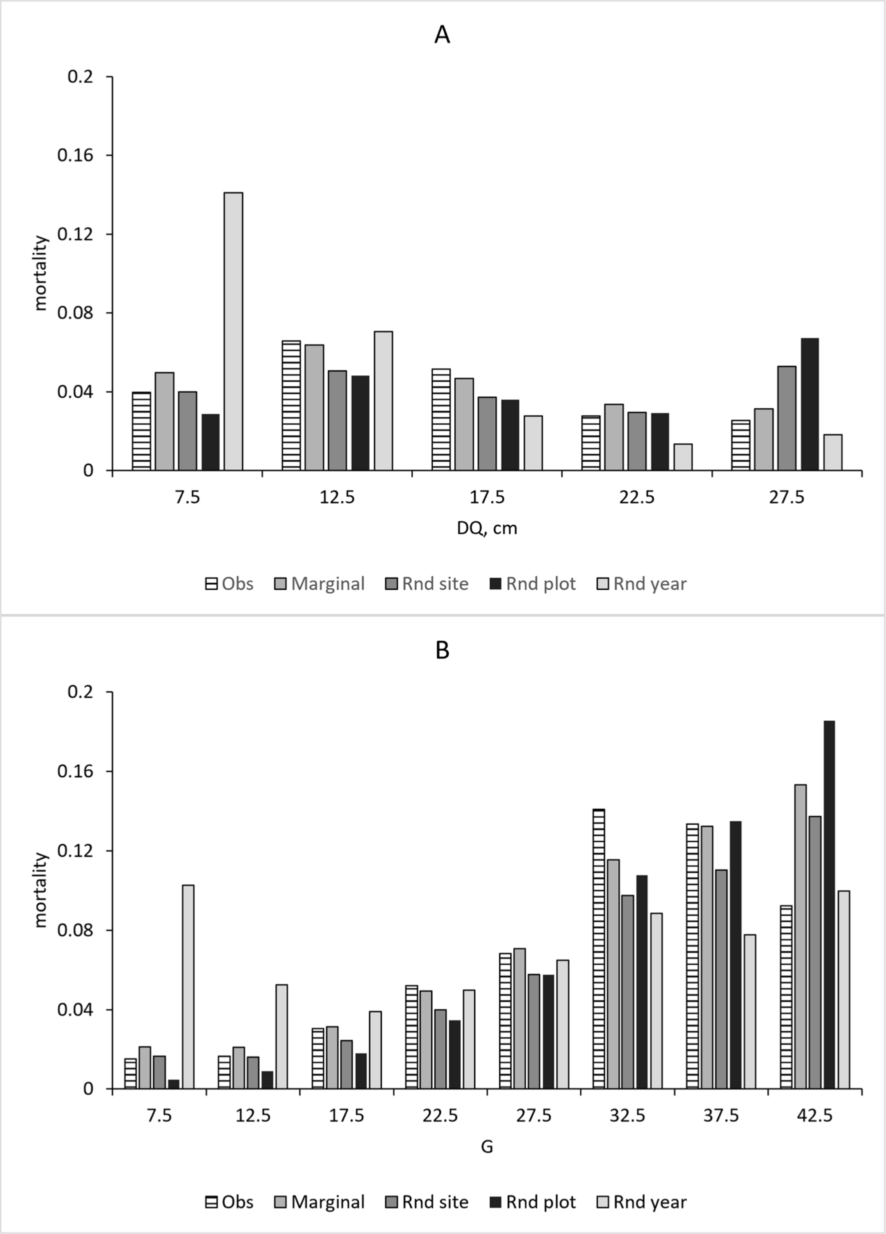

Fig. 4 The observed (obs) and predicted mortality rates with respect to quadratic mean diameter (*DQ*) and basal area (*G*) in the test dataset using the marginal and mixed-effects mortality models with random site, plot, and year effects. Note that the y-axis is capped at a 0.2 mortality rate

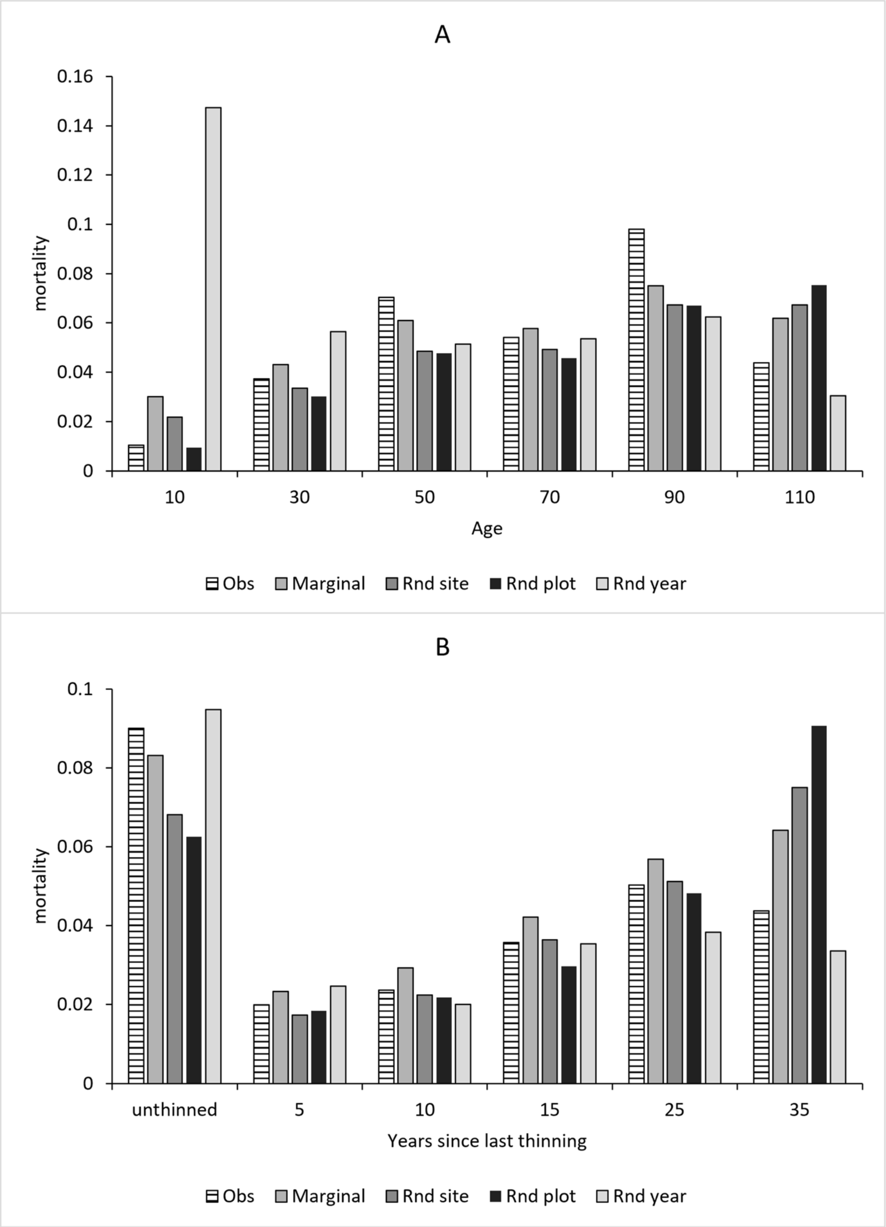

Fig. 5 The observed (obs) and predicted mortality rates with respect to stand age and years since last thinning in the test dataset using the marginal and mixed-effects mortality models with random site, plot, and year effects. Note that the y-axis is capped at a 0.16 mortality rate

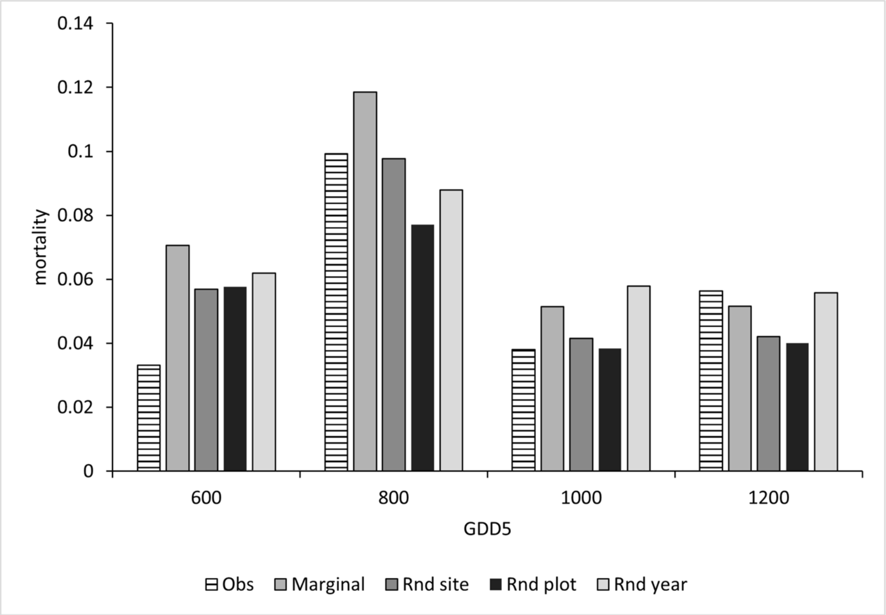

Fig. 6 Mortality rate in the test datasets with respect to growing degree days (*GDD5* as a 30-year average temperature sum from 1971 to 2000 with a 5 °C threshold) using the marginal and mixed-effects mortality models with random site, plot, and year effects

predictions from 600 to 1000 *GDD5* classes, but underestimated mortality in the 1200 *CDD5* class more than the marginal and random year models.

Motti Simulation with the Current and New Mortality Models

The mortality models produced satisfactory results when the 16 to 31 year simulations with the Motti simulator were compared with the observed stand density (Table 3). The marginal model demonstrated the most accurate for the development of stem number *N* (– 1.9% bias and 12.4% RMSE) and simultaneously for basal area *G*, with a

Table 3 The evaluation results of the Motti simulations in the 10 unthinned plots of the test dataset, with the marginal and mixed-effects logistic models applied with random site, plot, and year effects

Model	Marginal		Random site		Random plot(site)		Random year(site)	
Variable	*N*	*G*	*N*	*G*	*N*	*G*	*N*	*G*
Bias	**– 28.7**	**– 0.2**	– 65.7	– 0.8	– 53.3	– 0.6	– 81.2	– 1.2
Bias%	**– 1.9**	**– 0.7**	– 4.2	– 2.5	– 3.4	– 1.8	– 5.2	– 3.7
Stdev	**162.7**	**1.6**	170.7	1.8	165.6	1.7	212.4	2.1
RMSE	**191.6**	**1.9**	212.5	2.3	201.9	21	264.1	2.8
RMSE%	**12.4**	**6.0**	13.7	7.4	13.0	6.7	17.0	8.9

0.7% overestimation and 6.0% RMSE (Table 3). The mixed model with a random year effect produced the least accurate results for N and G, specifically, the bias was 5.2% and 3.7%, and the RMSE was 17.0% and 9.0%, respectively. The current model of the Motti simulator (Eq. 7) showed a 3.5% underestimation in N, with a standard deviation of the prediction error (stdev) of 220 and an RMSE of 12.7%. The corresponding figures for G were a 1.2% underestimation, 1.8 stdev and 5.0% RMSE. Therefor, the number of stems was marginally more biased when predicted with the current *RDFL*-based model than with the new *BAL*-based marginal model, but generally slightly better than with the new mixed models. Similarly, the predicted basal area was slightly less biased with the new marginal model. However, the current *RDFL*-based model provided a smaller bias in G than the new mixed models. Finally, the RMSE% for G with the current model was 5%, which was lower than with the new marginal model.

The new models are designed for use without a stand-level upper limit, based on the self-thinning line by Hynynen (1993). The mean of the predicted stand-level maximum basal areas were higher when using the new marginal model (45.5 m^2 ha^{-1}) and even higher with the random site effect model (46.0 m^2 ha^{-1}) compared to the current *RDFL*-based model (42.3 m^2 ha^{-1}) (Table 4). The random year effect model resulted in a much higher mean (68.1 m^2 ha^{-1}). It is worth noting that the highest observed basal area was 43.9 m^2 ha^{-1}, and that the trend was for increasing G for each case. This indicates that the absolute maximum was not yet reached. However, the mixed model with the random year effect reached an implausible maximum G on average, with the highest value reaching 81.4 m^2 ha^{-1}. This outcome was anticipated, given the illogical negative parameter for ln(G). Nevertheless, it was deemed appropriate to evaluate this model as an exemple of a model with the best statistical fit in comparison to models characterised by ecologically reasonable parameters (see Yang et al. 2003). Finally, the models, with the exception of the mixed model with a random year effect, can be used without an upper restriction based on the self-thinning line.

Discussion

The use of a relatively straightforward mortality model has proved advantageous for the logical model behaviour in terms of the estimated effect of the included variables and the predicted versus observed mortality rates against these variables (Jutras

Table 4 The maximum basal areas (G) predicted using the marginal and mixed-effects logistic models with random site, plot, and year effects, as well as the current *RDFL*-based model, in the prolonged Motti simulation, and the maximum observed G of the monitoring period

Maximum G	Marginal	Rnd site	Rnd plot	Rnd year	*RDFL*	Max obs
Average	45.5	46.0	42.1	68.1	42.3	37.3
Minimum	39.9	41.0	38.2	60.0	36.6	28.0
Maximum	51.8	54.5	48.5	81.4	45.9	43.9
Stdev	3.5	4.0	3.0	6.6	3.3	4.9

et al. 2003; Monserud and Sterba 1999; Monserud et al. 2005). We employed tree *dbh* to model increased mortality for small trees, as well as $dbh^{1.2}$ to account for slightly elevated mortality rates among the largest pine trees. In fact, dbh^2 and $dbh^{1.5}$ provided a slightly better fit statistic (e.g., 70 and 30 units smaller AIC), but then the predicted mortality rate for the largest trees became unrealistically high in each mixed model. It is notable that, in contrast to the final model behaviour illustrated in Fig. 2, the application of exponents 2 and 1.5 would result in the demise of the vast majority of trees with a diameter of 60 cm. The competition at the tree level was described by $BAL/\sqrt{(dbh)}$. The *BAL* was divided by $\sqrt{(dbh)}$ because the *BAL* alone in a model resulted in an overestimation of mortality for the highest *BAL* values. While the $BAL/\sqrt{(dbh)}$ was the essential driver in modelling the corresponding survival rate for Norway spruce (Siipilehto et al. 2021), it also proved to be the optimal choice when a single *BAL*-based variable was included in the pine model. It is optimal that the individual tree survival or mortality model can be used without self-thinning constraints (Monserud et al. 2005). In our study, all models except the random year model demonstrated a reasonable maximum basal area in the prolonged simulation run without thinning.

Age is seldom used to describe the developmental status of a stand in a mortality/survival model. However, recently some models have been developed to integrate the effects of competition and ageing on the survival probability of spruce trees in Finland and Norway. The results demonstrated that senescence was a significant factor in predicting spruce survival (Siipilehto et al. 2021). The stand age was found to have a significant impact on the survival probability of large spruce trees in both managed stands and old-growth stands (Siipilehto et al. 2021). As Scots pine is a long-lived species, reaching an age of over 500 years (even 700–800 years according to Engelmark and Hytteborn 1999), while the maximum age of Norway spruce is close to 300 years (Zackrisson et al. 1995; Kuuluvainen et al. 2002), we found that the age of pine was not a significant factor within the range of our datasets. Upon evaluation of the models in this study against age to ensure no bias was present in terms of age, it was found that age was not as significant a factor as it is for spruce. Nowadays, forest inventories frequently rely on airborne laser scanning (ALS) data, which enables more accurate estimation of stand density, especially the basal area, and mean dimensions (*DG* and *HG*) than the stand age (Maltamo et al. 2020). Furthermore, if a stand is managed according to uneven age management principles, the stand age is not a relevant variable (Moser 1972; Monserud and Sterba 1999; Peng 2000; Maltamo et al. 2020). We therefore chose to develop the models without including age, in order to ensure maximum flexibility for future applications. Instead of age, the developmental stage of a stand was defined by the *DQ*.

One might ask whether the potential impact of forest management should be included in the model. It is a common practice to fit survival or mortality models based on unthinned stands only, or at least to ignore the thinning effect (Yao et al. 2001; Temesgen and Mitchell 2005; Vieilledent et al. 2009). It should be noted that there are a few exceptions to this, namely the models by Fridman and Ståhl (2001), Sims et al. (2014) and Siipilehto et al. (2021), which include the effect of forest management (thinning) in their models. Alternatively, in models incorporating a crown ratio or diameter increment, the impact of forest management is incorporated indirectly through enhanced tree vitality (Hamilton 1986). As the mortality models in this study are designed for use in a stand simulator, it is crucial to consider the impact of stand management on tree mortality. Furthermore, the mortality models were developed based on experiments where a reduction of tree mortality was due to thinning from below. Accordingly, the thinning effect was incorporated into the models as two dummy variables, delineating the time since the last thinning up to 10 years. Scots pine, as a shade intolerant species, benefited more from thinnings than Norway spruce, for which the recent thinnings within 5 years was not significant (Siipilehto et al. 2021). On the other hand, the recent thinnings significantly reduced mortality in Norway spruce in Sweden, but were not significant for Scots pine (Fridman and Ståhl 2001). In the USA, thinning from below decreased mortality, while intensive thinning from above increased mortality (Powers et al. 2010).

The evaluation of the models can be carried out in several ways. In addition to the goodness-of-fit statistics and the observed versus predicted mortality rates with respect to tree-, stand- and site-level variables, we evaluated the behaviour of the models by simulating the stand development over time. This revealed important information on how the models worked. For instance, the mixed model with a random year effect produced the best fit statistics. However, a significant shortcoming was that this model did not impose any restrictions on stand basal area development, in contrast to the other models. It is notable that the average maximum G was 68 m^2 ha^{-1} in the prolonged Motti simulation, with the highest predicted G reaching 81 m^2 ha^{-1}. These values represent an extremely high stand density for a shade intolerant species such as Scots pine. It is therefore recommended that the model with the random year effect should not be used without an upper limit of self-thinning restriction. In contrast, the marginal model yielded a reasonable average maximum G of 46 m^2 ha^{-1} for the unthinned stands in the prolonged Motti simulation. The existing *RDFL*-based model provided a lower maximum basal area, namely 42 m^2 ha^{-1}, while underestimating G in the test dataset by only 1.2%. However, should these datasets be measured in the future, the predicted mortality by the current *RDFL*-based model may result in higher underestimations for basal area predictions.

The assessment of the mortality models with varying random effects in the Motti simulations demonstrated that the enhanced statistical goodness-of-fit of a model (as indicated by the $-$ 2 log likelihood and AIC) did not translate into enhanced simulation outcomes. We therefore recommend applying the marginal model, which offers the best development of the stem number and the stand basal area over time. The current mortality model, based on the self-thinning line (Hynynen 1993; Hynynen et al. 2002), also demonstrated excellent results. It is worth noting that the maximum basal area in the prolonged Motti simulations was lower with the current model than that of the best performing *BAL*-based models. The observed basal area development continued to show an increasing trend, which may result in a more pronounced underestimation of the predicted maximum basal area by the current *RDFL*-based model. It may be possible to enhance the model further by fitting a new self-thinning model with the most recent data, specifically the latest measurements from the same experiments. Finally, we presented the mixed models with random effects for the intercept only. It would be feasible to utilise alternative random effect combinations, including random slopes within the model. However, we do not anticipate that incorporating additional mixed effects combinations will result in any notable enhancements in a practical application, particularly when compared with the best-performing marginal model.

Understanding tree mortality is a prerequisite to avoid overestimating stand yield. Integrating growth models with the mortality models developed in this study reduces uncertainties in stand predictions and helps to make recommendations for sustainable forest management. The advantage is the detailed representation of both tree and stand dynamics, including tree mortality. This provides forest managers with robust predictions of future stand development.

Acknowledgements The study was conducted in the Natural Resources Institute Finland (Luke) and supported by a grant from the Research Council of Finland (No 347782). We are grateful to the field staff of Luke for the maintaining and measuring the experiments over the years.

Authors' Contribution Jouni Siipilehto: methodology, data analyses and model validation, data curation, writing.

Harri Mäkinen: methodology, data curation, writing.

Funding Open access funding provided by Natural Resources Institute Finland. The study was mainly funded by Natural Resources Institute Finland (Luke) and supported by a grant from the Research Council of Finland (No 347782).

Data Availability Data of this research can be available on reasonable request.

Declarations

Competing Interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahtikoski, A., J. Siipilehto, H. Salminen, M. Lehtonen, and J. Hynynen. 2018. Effect of stand structure and number of sample trees on optimal management for Scots pine: a model-based study. *Forests* 9(12):750, 15 p. https://doi.org/10.3390/f9120750.

Boeck, A., J. Dieler, P. Biber, H. Pretzsch, and D.P. Ankers. 2014. Predicting tree mortality for European beech in southern Germany using spatially explicit competition indices. *Forest Science* 60 (4): 613–622. https://doi.org/10.5849/forsci.12-133.

Cajander, A.K. 1926. The theory of forest types. *Acta Forestalia Fennica* 29. 108 p.

Coble, D.W., and Y.J. Lee. 2008. A new diameter distribution model for unmanaged slash pine plantations in East Texas. *Southern Journal of Applied Forestry* 32 (2): 89–94. https://doi.org/10.1093/sjaf/32.2.89.

Eid, T., and E. Tuhus. 2001. Models for individual tree mortality in Norway. *Forest Ecology and Management* 154: 69–84. https://doi.org/10.1016/S0378-1127(00)00634-4.

Engelmark, O., and H. Hytteborn. 1999. Coniferous forests. *Acta Phytogeographica Suecica* 84: 55–74. ISBN 91–72 1 0–084–2. https://www.diva-portal.org/smash/get/diva2:565333/FULLTEXT01.pdfIngvar#page=57.

Fortin, M., S. Bédard, J. DeBlois, and S. Meunier. 2008. Predicting individual tree mortality in northern hardwood stands under uneven-aged management in southern Québec Canada. *Annals of Forest Science* 65 (2008): 205. https://doi.org/10.1051/forest:2007088.

Fridman, J., and G. Ståhl. 2001. A three-step approach for modelling tree mortality in Swedish forests. *Scandinavian Journal of Forest Research* 16: 455–466. https://doi.org/10.1080/02827580152632856.

Goldstein, H. 1995. *Multilevel statistical models*. Kendall's Library of Statistics 3. Second Edition 178 p.

Hamilton, D.A. 1986. A logistic model for mortality in thinned and unthinned mixed conifer stands of northern Idaho. *Forest Science* 32: 989–1000. https://doi.org/10.1093/forestscience/32.4.989.

Hynynen, J. 1993. Self-thinning models for even-aged stands of *Pinus sylvestris*, *Picea abies* and *Betula pendula*. *Scandinavian Journal of Forest Research* 8 (3): 326–336. https://doi.org/10.1080/02827589309382781.

Hynynen, J., R. Ojansuu, H. Hökkä, J. Siipilehto, H. Salminen, and P. Haapala. 2002. *Models for predicting stand development in MELA system.* Finnish Forest Institute, Research Papers 835, Vantaa Research Center. 116 p. http://urn.fi/URN:ISBN:951-40-1815-X.

Hynynen, J., H. Salminen, A. Ahtikoski, S. Huuskonen, R. Ojansuu, J. Siipilehto, M. Lehtonen, A. Rummukainen, S. Kojola, and K. Eerikäinen. 2014. Scenario analysis for the biomass supply potential and the future development of Finnish forest resources. *Working Papers of the Finnish Forest Research Institute* 302. 106 p. http://urn.fi/URN:ISBN:978-951-40-2487-0.

Ilvessalo, Y. 1932. The establishment and measurement of permanent sample plots in Suomi (Finland). *Communicationes Instituti Forestalis Fenniae* 17(2): 1–39. http://urn.fi/URN:NBN:fi-metla-201206141012.

Jutras, S., H. Hökkä, V. Alenius, and H. Salminen. 2003. Modeling mortality of individual trees in drained peatland sites in Finland. *Silva Fennica* 37(2): 235–251. https://doi.org/10.14214/sf.504.

Kiernan, D., E. Bevilacqua, R. Nyland, and L. Zhang. 2009. Modeling tree mortality in low- to medium-density uneven-aged hardwood stands under a selection system using generalized estimating equations. *Forest Science* 55 (4): 343–351. https://doi.org/10.1093/forestscience/55.4.343.

Kuuluvainen, T., L. Karjalainen, and H. Lehtonen. 2002. The age distribution in old-growth forest sites in Vienansalo wilderness, eastern Fennoscandia. *Silva Fennica* 36(1): 169–184. https://doi.org/10.14214/sf.556.

Laarmann, D., H. Korjus, A. Sims, J.A. Stanturf, A. Kiviste, and K. Körster. 2009. Analysis of naturalness and tree mortality patterns in Estonia. *Forest Ecology and Management* 258: 187–195. https://doi.org/10.1016/j.foreco.2009.07.014.

Laasasenaho, J. 1975. *Runkopuun saannon riippuvuus kannon korkeudesta ja latvan katkaisuläpimitasta.* Folia Forestalia 233. 20 p. http://urn.fi/URN:ISBN:951-40-0164-8.

Ma, Z., C. Peng, W. Li, Q. Zhu, W. Wang, X. Song, and J. Liu. 2013. Modeling individual tree mortality rates using marginal and random effects regression models. *Natural Resource Modeling* 26 (2): 131–153. https://doi.org/10.1111/j.1939-7445.2012.00124.x.

Mäkinen, H., and A. Isomäki. 2004. Thinning intensity and growth of Scots pine stands in Finland. *Forest Ecology and Management* 201: 311–325. https://doi.org/10.1016/j.foreco.2004.07.016.

Maltamo, M., H. Kinnunen, A. Kangas, and L. Korhonen. 2020. Predicting stand age in managed forests using National Forest Inventory field data and airborne laser scanning. *Forest Ecosystems* 7(44). https://doi.org/10.1186/s40663-020-00254-z.

Mielikäinen, K. 1979. Alaharvennuksen vaikutus männikön tuotokseen ja arvoon. *Folia Forestalia* 401, 1–23. http://urn.fi/URN:ISBN:951-40-0405-1.

Monserud, R. 1976. Simulation of forest tree mortality. *Forest Science* 22: 438–444.

Monserud, R., and H. Sterba. 1999. Modelling individual tree mortality for Austrian forest species. *Forest Ecology and Management* 113: 109–123. https://doi.org/10.1016/S0378-1127(98)00419-8.

Monserud, R., T. Ledermann, and H. Sterba. 2005. Are self-thinning constraints needed in a tree-specific mortality model. *Forest Science* 50 (6): 848–858. https://doi.org/10.1093/forestscience/50.6.848.

Moser, J.W., Jr. 1972. Dynamics of an uneven-aged forest stand. *Forest Science* 18 (3): 184–191.

Myllymäki, M., M. Kuronen, S. Bianchi, A. Pommerening, and L. Mehtätalo. 2024. A Bayesian approach to projecting forest dynamics and related uncertainty: An application to continuous cover forests. *Ecological Modelling* 491: 1–14. https://doi.org/10.1016/j.ecolmodel.2024.110669.

Newton, P.F., Y. Lei, and S.Y. Zhang. 2005. Stand-level diameter distribution yield model for black spruce plantations. *Forest Ecology and Management* 209 (3): 181–192. https://doi.org/10.1016/j.foreco.2005.01.020.

Nyyssönen, A. 1950. Vertailevia havaintoja hoidettujen ja luonnontilaisten männiköiden rakenteesta. *Silva Fennica* 68: 1–48. https://doi.org/10.14214/sf.a9091.

Peng, C. 2000. Growth and yield models for uneven-aged stands: Past, present and future. *Forest Ecology and Management* 132 (2–3): 259–279. https://doi.org/10.1016/S0378-1127(99)00229-7.

Powers, M.D., B.J. Palik, J.B. Bradford, S. Fraver, and C.R. Webster. 2010. Thinning method and intensity influence long-term mortality trends in a red pine forest. *Forest Ecology and Management* 260 (7): 1138–1148. https://doi.org/10.1016/j.foreco.2010.07.002.

Reineke, L.H. 1933. Perfecting a stand-density index for even-aged forests. *Journal of Agricultural Research* 46: 627–638.

Salas-Eljatib, C., and A.R. Weiskittel. 2020. On studying the patterns of individual-based tree mortality in natural forests: A modelling analysis. *Forest Ecology and Management* 475:118369. https://doi.org/10.1016/j.foreco.2020.118369.

Salminen, H., M. Lehtonen, and J. Hynynen. 2005. Reusing legacy FORTRAN in MOTTI growth and yield simulator. *Computers and Electronics in Agriculture* 49 (1): 105–113. https://doi.org/10.1016/j.compag.2005.02.005.

SAS Institute Inc. (2017) *SAS/STAT® 14.3 User's Guide: The GLIMMIX Procedure*. SAS Institute Inc., Cary, NC, USA.

Siipilehto, J., and A. Kangas. 2015. Näslundin pituuskäyrä ja siihen perustuvia malleja läpimitta-pituus riippuvuudesta suomalaisissa talousmetsissä. *Metsätieteen aikakauskirja* 4/2015. http://urn.fi/URN:NBN:fi-fe2016102425504.

Siipilehto, J., and L. Mehtätalo. 2013. Parameter recovery vs. parameter prediction for the Weibull distribution validated for Scots pine stands in Finland. *Silva Fennica* 47(4): 1–22. http://www.silvafennica.fi/article/1057.

Siipilehto, J., H. Mäkinen, K. Andreassen, and M. Peltoniemi. 2021. Models for integrating and identifying the effect of senescence on individual tree survival probability for Norway spruce. *Silva Fennica* 55(2): 10496. https://doi.org/10.14214/sf.10496.

Sims, A., R. Mändma, D. Laarmann, and H. Korjus. 2014. Assessment of tree mortality on Estonian Network of Forest Research Plots. *Forestry Studies | Metsanduslikud Uurimused* 60: 57−68. https://doi.org/10.2478/fsmu-2014-0005.

Stiratelli, R., N. Laird, and J.H. Ware. 1984. Random-effects models for serial observations with binary response. *Biometrics* 961–971. https://doi.org/10.2307/2531147.

Temesgen, H., and S.J. Mitchell. 2005. An individual-tree mortality model for complex stands of southeastern British Columbia. *Western Journal of Applied Forestry* 20 (2): 101–109. https://doi.org/10.1093/wjaf/20.2.101.

Vanclay, J. 1994. *Modelling forest growth and yield applications to mixed tropical forests*. CAB International. 312 p. ISBN 0 85198 913 6.

Vieilledent, G., B. Courbaud, G. Kunstler, J.F. Dhôte, and J.S. Clark. 2009. Biases in the estimation of size-dependent mortality models: Advantages of a semiparametric approach. *Canadian Journal of Forest Research* 39: 1430–1443. https://doi.org/10.1139/x09-047.

Vuokila, Y. 1962. The effect of thinnings on the yield of pine and birch stands. *Communicationes Instituti Forestalis Fenniae* 55: 12 pp.

Weiskittel, A., P. Gould, and H. Temesgen. 2009. Sources of variation in the self-thinning boundary line for three species with varying levels of shade tolerance. *Forest Science* 55 (1): 84–93. https://doi.org/10.1093/forestscience/55.1.84.

Wykoff, W.R. 1990. A basal area increment model for individual conifers in the northern Rocky Mountains. *Forest Science* 36 (4): 1077–1104. https://doi.org/10.1093/forestscience/36.4.1077.

Yang, Y., and S. Huang. 2013. A generalized mixed logistic model for predicting individual tree survival probability with unequal measurement intervals. *Forest Science* 59: 177–187. https://doi.org/10.5849/forsci.10-092.

Yang, Y., S.J. Titus, and S. Huang. 2003. Modelling individual tree mortality for white pine in Alberta. *Ecological Modelling* 163: 209–222. https://doi.org/10.1016/S0304-3800(03)00008-5.

Yao, X., S.J. Titus, and S.E. MacDonald. 2001. A generalized logistic model of individual tree mortality for aspen, white spruce, and lodgepole pine in Alberta mixedwood forests. *Canadian Journal of Forest Research* 31: 283–291. https://doi.org/10.1139/x00-162.

Yli-Kojola, H. 2005. *Metsikkö- ja puutuhojen ennustemallit*. Metsäntutkimuslaitoksen tiedonantoja 948. http://urn.fi/URN:ISBN:951-40-1989-X.

Zackrisson, O., M.C. Steijlen, and G. Hörnberg. 1995. Regeneration pulses and climate-vegetation interactions in nonpyrogenic boreal Scots pine stands. *Journal of Ecology* 83: 1–15. https://doi.org/10.2307/2261600.

Authors and Affiliations

Jouni Siipilehto[1] · **Harri Mäkinen**[1]

✉ Harri Mäkinen
harri.makinen@luke.fi

Jouni Siipilehto
jouni.siipilehto@luke.fi

[1] Natural Resources Institute Finland (Luke), Latokartanonkaari 9, 00790 Helsinki, Finland