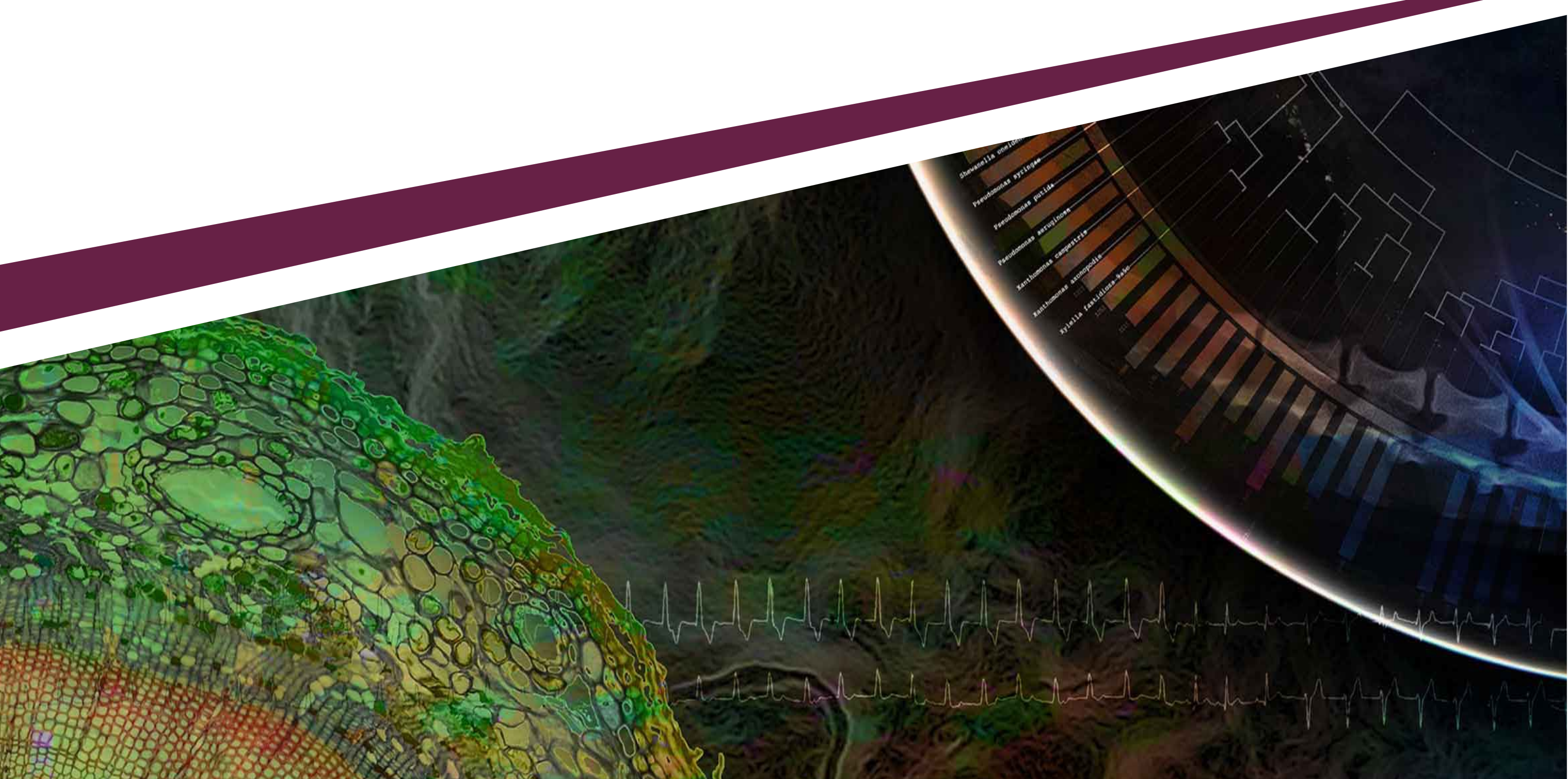
Basal area growth and damage prediction models for retained pine seed trees in Sweden

Alex Appiah Mensah

Swedish University of Agricultural Sciences, SLU
Department of Forest Resource Management

Arbetsrapport / Sveriges lantbruksuniversitet, Institutionen för skoglig resurshushållning, 557
ISSN: 1401-1204
2024



Basal area growth and damage prediction models for retained pine seed trees in Sweden

Alex Appiah Mensah, https://orcid.org/0000-0001-5073-422X, Swedish University of Agricultural Sciences, Department of Forest Resource Management

Publisher:	Swedish University of Agricultural Sciences, Department of Forest Resource Management
Year of publication:	2024
Place of publication:	**Umeå**
Copyright:	All featured images are used with permission from the copyright owner.
Title of series:	Arbetsrapport / Sveriges lantbruksuniversitet, Institutionen för skoglig resurshushållning
Part number:	557
ISSN:	1401-1204
Keywords:	Forest management planning, forest yield research, Scots pine, growth, damage

Abstract

Natural pine regeneration with seed trees is an established and well-used method in Swedish forestry. Retained pines respond to increased exposure and resource availability by increasing their radial stem growth, as observed in the annual rings' width after release. Currently, the basal area growth of seed trees left over after clear felling is simulated (in Heureka) with individual tree models dependent on the tree, competition and site factors. However, a major drawback of those models is that the growth reaction of retained seed trees after liberation is not accounted for leading to potentially biased estimates of the expected growth. This report presents functions for predicting the basal area growth underbark (BAI) for retained pine seed trees. In parallel, models for damage predictions are provided.

The basal growth functions were developed in two steps. First, the response after liberation was examined and after that quantified using a logistic growth function. The results showed a significant increase in the BAI, which persisted up to 30 years after liberation and the magnitude of the response depended on the site index. On sites with high site index, the relative increase in growth 10 years before and after liberation was 11%. The corresponding value for sites with low site index was 21%. The maximum relative response culminated at six and seven years after liberation for high and low site indices, respectively. In the second step, the liberation response model was integrated with a BAI prediction model with the predictor variables diameter at breast height, crown ratio, site index and categorical indicators representing stand management class (from barelands to unthinned stands). The new function's simulated outcome demonstrates logical predictions of growth.

The regression function for predicting the probability of retained pine trees being damaged at a given time was calibrated with tree height, age and site index as predictor variables. The model accuracy using a cut-off value of 0.5 was moderate (i.e., 66%) and with sensitivity and specificity values of 80% and 40%, respectively. Data from the temporary sample plots of the Swedish NFI distributed throughout the country facilitated this investigation. The limitations and potential developments of the growth and damage models are discussed.

Keywords: Forest management planning, forest yield research, Scots pine, growth, damage.

Sammanfattning

Naturlig tallföryngring med fröträd är en etablerad och välanvänd metod i svenskt skogsbruk. Kvarlämnade tallar svarar på ökad exponering och resurstillgång genom att öka sin radiella stamtillväxt, vilket observeras i årsringarnas bredd efter friställning. För närvarande simuleras grundytetillväxten hos fröträd som lämnats kvar efter kalavverkning (i Heureka) med individuella trädmodeller som beror av träd-, konkurrens- och ståndortsfaktorer. En stor nackdel med dessa modeller är dock att de kvarlämnade fröträdens tillväxtreaktion efter friställning inte beaktas, vilket leder till potentiellt felaktiga skattningar av den förväntade tillväxten. I denna rapport presenteras funktioner för att prediktera tillväxten av grundyta under bark (BAI) för kvarlämnade tallfröträd. Parallellt med detta presenteras modeller för skadeprognoser.

Funktionerna för grundytetillväxt utvecklades i två steg. Först undersöktes responsen efter friställningen och därefter kvantifierades den med hjälp av en logistisk tillväxtfunktion. Resultaten visade en signifikant ökning av BAI, som kvarstod upp till 30 år efter friställningen, och storleken på responsen berodde av ståndortsindex. På platser med högt ståndortsindex var den relativa ökningen av tillväxten 11%, 10 år efter friställningen. Motsvarande värde för platser med lågt ståndortsindex var 21%. Den maximala relativa responsen kulminerade sex respektive sju år efter friställningen för höga respektive låga ståndortsindex. I det andra steget integrerades modellen för friställning med en BAI-prediktionsmodell med oberoende variabler diameter i brösthöjd, kronandel, ståndortsindex och kategoriska indikatorer som representerar beståndets skötselklass (tex barmark till ogallrade bestånd). Den nya funktionens simulerade utfall visar på logiska förutsägelser om tillväxt.

Regressionsfunktionen för att förutsäga sannolikheten för att kvarlämnade tallar skadas en viss tidpunkt efter friställning kalibrerades med trädhöjd, ålder och ståndortsindex som oberoende variabler. Modellens noggrannhet med ett gränsvärde på 0,5 var måttlig (dvs. 66%) och med känslighets- och specificitetsvärden på 80% respektive 40%. Denna undersökning baserades på data från Riksskogstaxeringens tillfälliga provytor, vilka representerar hela landet. Tillväxt- och skademodellernas potentiella begränsningar och noggrannhet av framskrivningar diskuteras.

Nyckelord: Planering av skogsbruk, forskning om skogsavkastning, tall, tillväxt, skador.

Table of contents

1. Introduction

Natural regeneration forms an integral part of Swedish forestry. Over the last 40 years, about 200000 hectares of forestland have been regenerated annually and out of this, about 30000 – 60000 constitute natural regeneration (Karlsson et al., 2017). Most natural regeneration is carried out as seed trees in pine forests (*Pinus sylvestris* L.), and of all pine regenerations, the proportional use of the seed tree method ranges between 30 – 60%. The corresponding figure for Norway spruce is about 1%. Thus, pine is better suited to natural regeneration with seed trees than spruce (National Board of Forestry, 1998). The principal reasons are that pine (1) has a more even seed production, (2) is less windthrown, (3) there is less risk of damage and the seed trees grow more economically, and (4) soil preparation is easier to perform under pine seed stands. Near-term spruce regeneration can take significantly longer (Karlsson et al., 2017). From bud to seed dispersal takes three years (four calendar years) in pine and 1.5-2 years (2-3 calendar years) in spruce. High temperatures, good light and nutrient availability favour seed production. Large trees generally produce more seeds than smaller ones (Karlsson et al., 2017).

Generally, for this regeneration strategy, densities of 50 – 150 seed trees ha^{-1} with basal areas of 5 – 10 m^2 ha^{-1} and uniformly distributed throughout the regeneration area are usually recommended in Sweden (Karlsson and Örlander, 2004). When the density of seed trees increases, the regeneration strategy becomes more or less similar to the "shelterwood method" (Shultz, 1997). Seed trees have several advantages, such as cheaper regeneration costs by providing freely available seeds and seedlings, maintenance of genetic make-ups of the former stand, and sheltering new seedlings. Additionally, removing seed trees from the new stand during first thinning operations may provide some extra income. Denser seed trees reduce the risk of frost and insect damage to plants, keep down competing vegetation, mitigate rising groundwater levels and reduce nitrogen leaching after harvesting. They also benefit biodiversity and the landscape. However, the overstory trees must be removed in time so as not to inhibit regeneration (National Board of Forestry, 1998).

1.1 Growth prediction in the Heureka DSS

Retained pines are defined here as trees deliberately omitted from harvests. Like trees growing in thinned stands, liberated trees also respond to increased exposure and resource availability by increasing their radial stem growth, as observed in the annual rings' width, crown and root growth. There are examples of trees grown in dense stands that increased their ring width by a factor of four over 3 – 5 years after release by thinning. As measured at breast height, the response is sometimes delayed. For Norway spruce, a delay of two to four years depending on the site type is reported (Näslund, 1942). Valinger (1992) found an immediate response in the lower stem parts after the first thinning for Scots pine. The duration of the increased increment after the liberation has been reported as 15-25 years for Scots pine in Sweden (Hagberg, 1942; Jonsson, 1995). The growth reaction is expected to decrease as competition with and shelter by the new stand increases (Jacobsson, 2004).

Currently, in Heureka, the basal area growth of seed trees left over after clear felling is simulated with individual tree models dependent on the tree size dimensions, competition and site factors (Elfving, 2010). The models were calibrated with data from repeated measurements on permanent plots of the Swedish National Forestry (NFI), estimating the 5-year basal area increments. However, a major drawback of those models is that the growth reaction of retained seed trees after liberation (i.e., after clear-cut) is not implemented leading to potentially biased estimates of the expected growth.

Now, abundant information on stem radial growth from increment cores of retained pine seed trees is available from the NFI allowing for a thorough assessment of the growth response after liberation. Further, the data offers the possibility to develop robust functions to predict basal area growth in annual time steps (i.e., annual increments) or periods (i.e. periodic growth such as 5-year periods). The data also contains information about damages offering the possibility to explore and develop predictive models of the damage risk of pine seed trees.

1.2 Research objectives

This project focuses on developing models for annual basal area increment underbark and damage risk for retained pine seed trees. The approach for the growth modelling is as follows: (1) examination of the growth reaction after liberation and (2) general growth model calibrated with the release response from step 1. Radial growth and other mensurational data from the temporary plots oif the Swedish NFI was used for the investigation.

2. Materials and methods

2.1 Data

The Swedish landscape is dominated by temperate and boreal forest conditions between latitudes 55.4 and 68.4 °N and altitudes from sea level to the arctic tree line, 500–800 m above the sea level. The large gradient in forest growth offers the possibility of evaluating growth responses at different sites. The Swedish NFI data form an ideal database for growth studies. Representative data on stands and trees have been gathered since 1923 with the methods of measurement being fixed for long periods. The Swedish NFI is a probability sample characterised by a systematic cluster sampling design (Fridman et al., 2014). The design is adapted for the different parts (i.e., mainly through five inventory regions) of the country and covers large gradients in management and climatic conditions. Annual measurements are conducted through a set of permanent and temporary sample plots. The permanent plots (radius 10 m) were established between 1983 and 1987 (about 20,000 plots on forestlands) and re-measured for the first time after five years (i.e., between 1988 and 1992). Temporary plots (radius 7 m) have been used since 1923 up to the present, and they provide radial growth data from increment cores taken from sample trees.

The annual resolution of radial growth data provides a suitable framework for the assessment of diameter or basal area growth and response to liberation (i.e., release cutting). In this work, the extensive stem radial growth data obtained from increment cores of sample trees on temporary plots on all forestlands was used for the investigation. Plots were selected based certain criteria: for example, from remaining trees on recently harvested stands to established stands with retained older pines from the former stand. On each plot, the following information was retrieved for the trees (e.g., total height, diameter at breast height, crown characteristics), stand (e.g., age, density, the type and time of forest management activities), and site (e.g., altitude, distance to coast, ground vegetation, soil moisture and type, site fertility). The sample plots (347 plots) have been inventoried from 2004 to 2022. The spatial distribution of the plots is shown in (Figure 1). The summary of the data is shown in Table 1.

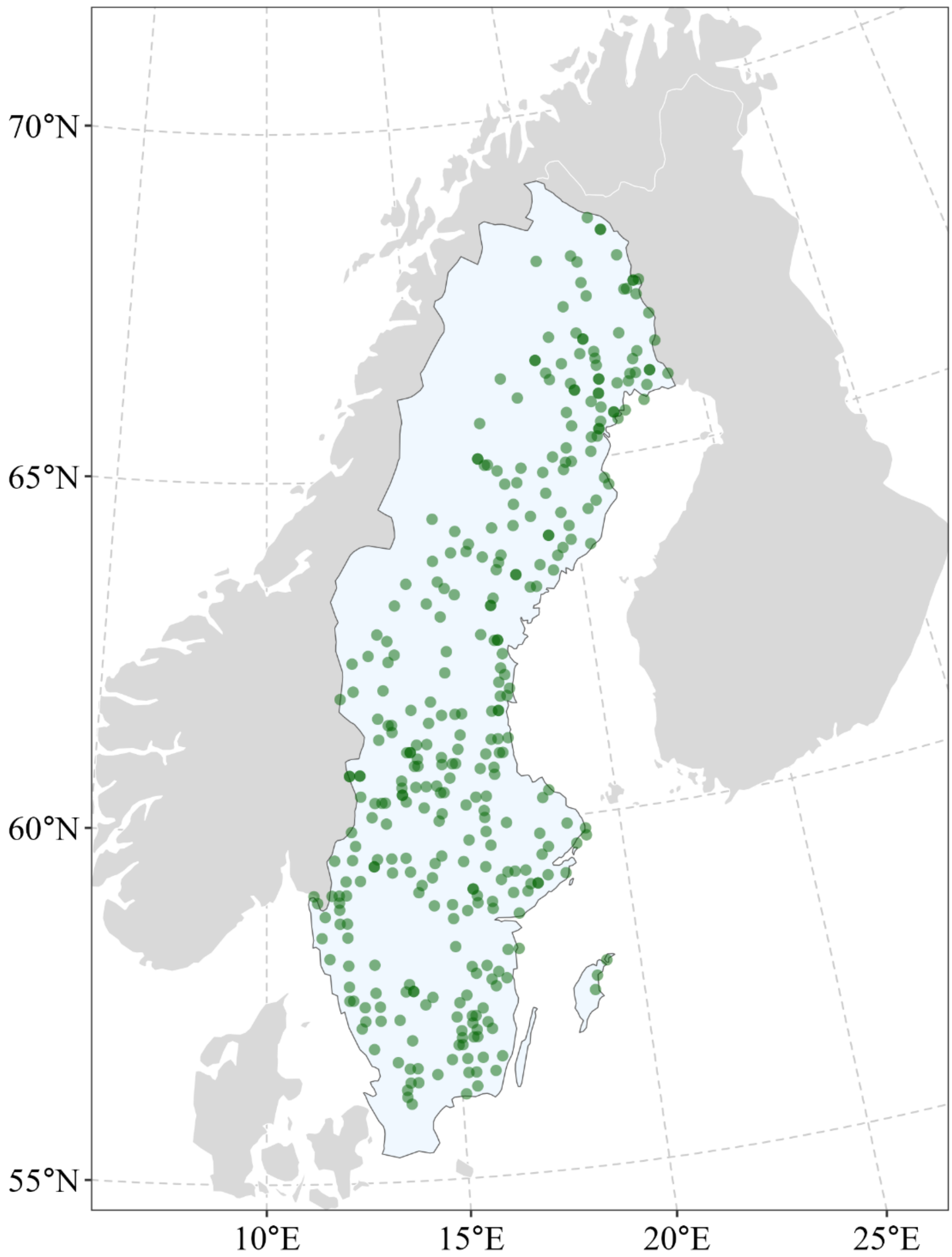


Figure 1. Locations of the NFI plots.

Table 1. Summary of field data.

Variable	Resolution	Min	Mean	Max	SD
Diameter, cm	Tree	6.0	28.6	34.41	9.37
BAI, cm^2/yr	Tree	2.75	14.52	45.72	9.22
Age, years	Tree	11	116	317	47
Site index, m	Stand	12	21.61	30	4.99
Stand height, m	Stand	0.10	4.21	22	4.56
Tree height, m	Tree	5.20	18.91	29.8	4.85

2.2 Methods

2.2.1 Modelling growth response after liberation

To predict the basal area growth dynamics of retained pine seed trees, it is necessary to investigate the magnitude of the growth response of the retained trees after liberation (i.e., release cutting). About 795 trees from 347 sample plots distributed throughout the country were analysed for the response after liberation. The retained pines had been liberated at different times. The exact times of the liberation were not known for all trees but were estimated from the age of the new stand and patterns of ring widths of the increment cores following Jakobsson (2004). As a prerequisite for minimising the variation in the growth response after liberation, the investigation was carried out for periods 10 and 20 years respectively, before and after the liberation.

The model to quantify the growth response after liberation was developed using annual basal area increment observations from liberation to 20 years after release. Preliminary tests indicated a culmination of the liberation response and the levels of maximum response in growth differed with site index. Thus, a bounded model (i.e., asymptotic) was used for the response analyses. Due to the differences in increment at the time of liberation (i.e., year 0) for each tree, the growth response at time t after liberation was expressed as a relative measure defined as the relative increase in basal area growth after liberation as compared to the growth at the time of liberation. This relative measure suggests that the response values are bounded between 0 and 1, with higher values indicating a larger growth reaction.

The expected basal area growth response of a tree at time t after liberation [E(Y|t)] was assumed to follow a three-parameter logistic model (Equation 1). The implicit form of the model was given as:

$$E(Y|t) = \frac{\alpha_1 SIS^{\alpha_2}}{1 + \exp^{(-\alpha_3(t-\alpha_4))}} \qquad (1)$$

where $\alpha_1 SIS^{\alpha_2}$ describes the asymptote (the limiting value when $t \to \infty$) which is modified by site index (SIS, site index according to site properties), α_3 measures the steepness of the curve and α_4 time-value at which the response reaches half of its asymptotic response.

The advantage of the logistic model is that the lower and maximum response predictions are bounded between 0 and 1 and thus, non-negative predicted values are avoided. The parameters α_1, α_2, α_3 and α_4 of Eq. 1 are unknown but are to be estimated from the given sample data (i.e., the NFI data). The parameters were estimated via a generalised nonlinear least squares procedure (using the gnls function from the nlme package in R). In addition, preliminary checks of the residual errors of (Eq. 1) indicated a heteroscedastic pattern as typical of longitudinal growth observations. Hence, a nonlinear variance function (constant plus power function of the predicted response - $\hat{y}$) was used to calibrate the residual variance of the model (Eq. 2) to achieve reliable and efficient estimates of parameters and associated standard errors.

$$V(\varepsilon) = \omega^2\left(\delta_0 + |\hat{y}|^{\delta_1}\right)^2 \qquad (2)$$

A similar model version was fitted with a first-order autocorrelation function but a likelihood ratio-test suggested no extra gain in information compared to a model without autocorrelation. Hence, the correlation parameter was excluded from the final model. A t-test analysis was carried out to examine the significance of the estimated model parameters at a 5% confidence level (i.e., whether the parameter estimates are statistically different from zero). All analyses were conducted using the R-Statistical Computing Environment (R Core Team, 2022).

2.2.2 Modelling basal area increment under the bark

Tree growth, for example, basal area increment can be perceived as the sum or product of the set of relevant growth factors. Thus, growth can be decomposed into several components, broadly categorised as a stand or climate-driven process (Zeide, 1993). The extent and magnitude by which tree ring widths capture interannual processes depend largely on the tree species and its surrounding environments (LeBlanc, 1990). An earlier study by Mensah et al. (2023) indicates that tree, stand and site variables explain about 52% of the variance in basal area growth.

To compute the annual basal area increment under bark (denoted as BAI), the function by Mensah et al. (2023) was used (see Equation 3 below). Typical input variables in the BAI function are the double ring width of the last measured full ring (denoted as DI), breast height diameter over bark at the time of inventory (dbh), and bark thickness. Bark thickness was estimated using the available functions for each species (Söderberg, 1992).

$$\mathrm{BAI} = (\mathrm{DUB}^2 - (\mathrm{DUB} - \mathrm{DI})^2) \times \frac{\pi}{4} \qquad (3)$$

where DUB is the diameter under the bark. Others are already defined.

Using the last measured full ring means that the growth associated with the inventory year 2021 refers to the growth year 2020. The annual ring for the inventory year was excluded since the inventory was performed during the growing season. On temporary plots, temporal variations in stand conditions are not readily discernible given that plots are measured only once in contrast to repeated measurements of permanent plots over time. Since growth is influenced by stand and site conditions over time, the use of the last growth ring means that growth can be linked to the stand conditions at the time of the inventory which may increase the variance explained by the estimated basal area increment model.

The basal area increment model presented by Mensah et al. (2023) provided a model basis for this study. The model predicting annual basal area increment under the bark was constructed as a function of tree dimensions, stand density, site fertility, and stand management. The tree dimensions comprised diameter at breast height, total tree height, and crown ratio. Stand density is expressed by basal area measured with a relascope from the centre of the plot. Site fertility was described by SIS (Hägglund and Ludmark, 1977). Management was characterised by the stage of stand development at the time of the inventory and comprised mainly of dummy variables describing barelands, seedlings, young stands under and above 3 m, and unthinned stands.

Mathematically, the function relating the annual basal area increment to a set of independent variables was expressed implicitly as:

$$\mathrm{BAI} = \mathrm{f}(\mathrm{dbh}, \mathrm{H}, \mathrm{Dens}, \mathrm{CR}, \mathrm{SIS}, \mathrm{H_g}, \mathrm{C_{ind}}, \mathrm{Y_{ind}}) + \varepsilon \quad (4)$$

where

BAI	= One-year increment of tree basal area under bark, cm^2 (estimated using Eq. 3).
dbh	= Tree diameter at breast height, over bark, at the time of inventory, cm.
H	= Total tree height at the time of inventory, m.
CR	= Crown ratio, defined as the ratio of the length of live crown to total tree height.
Dens	= Relative density describing stand closure (estimated from the basal area and mean height of the stand).
SIS	= Site index according to site properties, m (expresses the expected dominant height at 100 years).
H_g	= Basal area-weighted mean stand height, m.
C_{ind}	= Indicator for unthinned stands (1) and otherwise (0).
Y_{ind}	= Indicator for young stands above 3 m (1) and otherwise (0).

The dependent variable (BAI) was expressed in the logarithmic form to satisfy normality, additivity and variance homogeneity assumptions. Thus, an additive error term was considered. Appropriate transformations of the other numerical covariates were made to ensure residual homoscedasticity. Interaction terms were also defined among the predictor variables. The parameters of the log-linear basal area increment function f were estimated using the general multiple linear regression and the ordinary least squares method.

The model $\boldsymbol{f}$ was used to estimate the BAI in its original unit of measurement. The Baskerville's bias-correction procedure (Baskerville 1972) was applied to eliminate the problems of back-transformation bias as:

$$\mathrm{BAI} = \exp[\mathrm{f}(\ )] * \exp\left(\frac{\widehat{\sigma}^2}{2}\right) \quad (5)$$

2.2.3 Modelling tree-level damage probability

Tree and stand damage are observed by the NFI depending on the conditions of the living stands and trees. For stand damage, the assessment is carried out on a 20-m radius plot and the inventory method used varies for established (more than 7 m in average height) and young stands. In established stands, the assessment of damage refers only to trees in the following tree classes: independent, dominant, and co-dominant. In young stands, the assessment refers to main plants or stems. For tree damage, the assessment is made on the sample trees in either 7-m or 10-m radius plots, depending on whether the plot is permanent or temporary.

The following damages are reported: root, cambial, stem, and canopy damage. In addition, the dominant cause of damage to the living stand is given and includes factors of climate (snow, wind, frost, etc.), human (forest, twisted roots or other planting damage, etc.), vertebrate (moose, reindeer, wild boar, beaver, other larger mammals, etc.), insects (spruce bark beetle, etc.), fungus, fire, and others. The time of damage is generally reported for damages incurred during seasons 0-5 as well as damage of a continuous type.

In this study, only the damage of the sample trees was investigated on the temporary plots of the NFI. For seed trees, the most frequent damage was those from the factors of climate (e.g., wind and snow) and forestry (especially, root damage during harvest). Others were less frequent. Such unbalanced sample size was deemed not suitable to develop specific-damage models, but to explore a model for general damage prediction. Therefore, the analysis was made with damage status of a tree reclassified as damaged or undamaged. Only the observations with damages assumed to have occurred at the time of inventory (season 0) or previous season (season 1) or both (i.e., both previous and current season season) were used in the analysis in order to be able to pair the observed damage to the stand conditions at the time of inventory. This selection resulted in about 330 sample tree observations that were used to fit the model and the probability of damage is predicted on an annual time step.

The presence of damage on a tree in a plot over an interval can be modelled using logistic regression. The logistic function has been extensively used in the modelling of damage and mortality of trees and stands worldwide and has been shown to give satisfactory outcomes. Let y_{it} be the binary responses (where $y_{it} = 1$ when a tree was damaged at time t and $y_{it} = 0$ when the tree was not damaged) for observations i at time t, x_{it} include the predictors of that observation at measurement occasion t and β the regression coefficients, the y_{is} (dropping the subscript i for simplicity) is Bernoulli distributed with a mean μ (Eq. 6.1) and the μ is restricted to the range (0, 1) by a link function g. For binary data, a canonical logit function links the expected mean response linearly to the predictor variables (Eq. 6.2). Subsequently, the logistic model for predicting tree damage probability is given as (Eq. 6.3):

$$y_i = \text{Bernoulli}(\mu_i) \quad (6.1)$$

$$g(\mu) = \ln \frac{\mu}{1-\mu} \quad (6.2)$$

$$\pi_{it} = \frac{e^{x_{it}\beta}}{1+e^{x_{it}\beta}} \quad (6.3)$$

where in Eq. 6.3, π_{it} is the probability of damage occurring on tree i at measurement occasion t [i.e., $\Pr(y_{it} = 1)$], e is the base of the natural logarithm and β is a vector of parameters to be estimated from the sample data using maximum likelihood procedures. The predictors (x_s) tested comprised many variables such as those defined in section 2.2.2 and age at breast height.

3. Results

3.1 Growth response after liberation

Figure 2 shows the average annual basal area growth for trees larger than 10 cm at breast height before and after liberation. The magnitude of the growth response varied with the site index. On the sites with high site index, the absolute basal area increment 10 years before liberation fluctuated around 12 cm^2 and increased markedly after release, peaking at about 22 cm^2. The corresponding responses were 8.5 cm^2 and 12.4 cm^2 for sites with lower site index (Figure 2A). The duration of the increased increment continued up to 20 years after liberation. When comparing the increase in increment at any time after liberation to the growth at the time of liberation (i.e., year zero), the relative response increase culminated by a factor of 1.8 units at the seventh year after liberation on the lower site index. About 1.7 units in the sixth year after liberation were observed on the sites with high site index (Figure 2B).

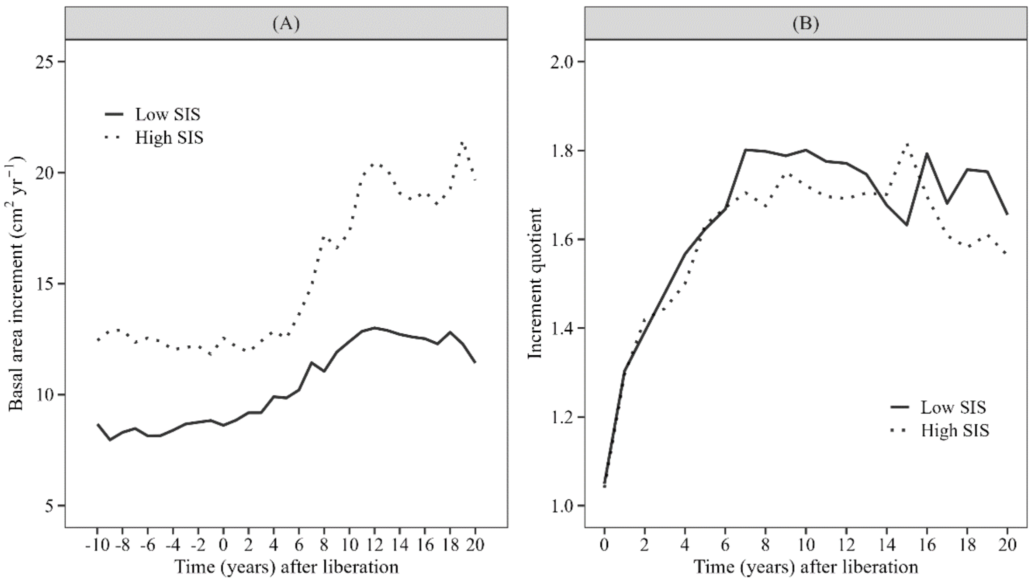


Figure 2. Average annual basal area growth under bark for trees larger than 10 cm at breast height during 10 years before and 20 years after liberation (A). Increment quotient of average annual basal area growth to basal area growth at

the year of liberation. The high site index refers to SIS > 22 m and the low site index refers to SIS ≤ 22 m.

The magnitude of the relative increase in growth 10 years before and after liberation was 11% on the high site index and 21% on the low site index (Table 2).

Table 2. Annual basal area increment (cm^2) at breast height for the 10 years before and after liberation separated by high and low site index.

Site index*	n	Before liberation	After liberation	Relative change (%)
High	2331	12.37±7.89	13.71±9.13	10.85
Low	2560	8.44±6.22	10.22±7.45	21.03

Data are means ± SD. * high site index refers to SIS > 22 m and low site index refers to SIS ≤ 22 m; n refers to the number of growth observations.

Table 3 shows the parameter estimates of the growth response model. Except for the parameter α_3, all others were significant at a 5% confidence level. The residual variability around the model (root mean square error RMSE) was 0.276 cm^2/year. The distribution of fitted residuals was homogeneous for the predicted response and time after liberation (Figure 3).

Table 3. Parameter estimates of the growth response model after liberation (Equation 1). SE is the standard error. CI is an asymptotic confidence interval of 95%.

Parameters	Estimate	SE	p-value	Lower CI (2.5%)	Upper CI (97.5%)
$\hat{\alpha}_1$	1.8853	0.4878	0.0001	0.9292	2.8414
$\hat{\alpha}_2$	-0.3983	0.0841	0.0000	-0.5631	-0.2335
$\hat{\alpha}_3$	-0.4106	0.6771	0.5444	0.1694	0.7724
$\hat{\alpha}_4$	0.4709	0.1538	0.0023	-1.7376	0.9164
	Residual error variance ($\hat{V}_\varepsilon$)				
$\hat{\omega}$	0.2319				
$\widehat{\delta_1}$	0.0054				
$\widehat{\delta_0}$	0.1002				
	Correlation between parameters				
$(\hat{\alpha}_1, \hat{\alpha}_2)$	-0.992				
$(\hat{\alpha}_1, \hat{\alpha}_3)$	-0.124				
$(\hat{\alpha}_1, \hat{\alpha}_4)$	-0.045				
$(\hat{\alpha}_2, \hat{\alpha}_3)$	0.037				
$(\hat{\alpha}_2, \hat{\alpha}_4)$	0.008				
$(\hat{\alpha}_3, \hat{\alpha}_4)$	0.816				

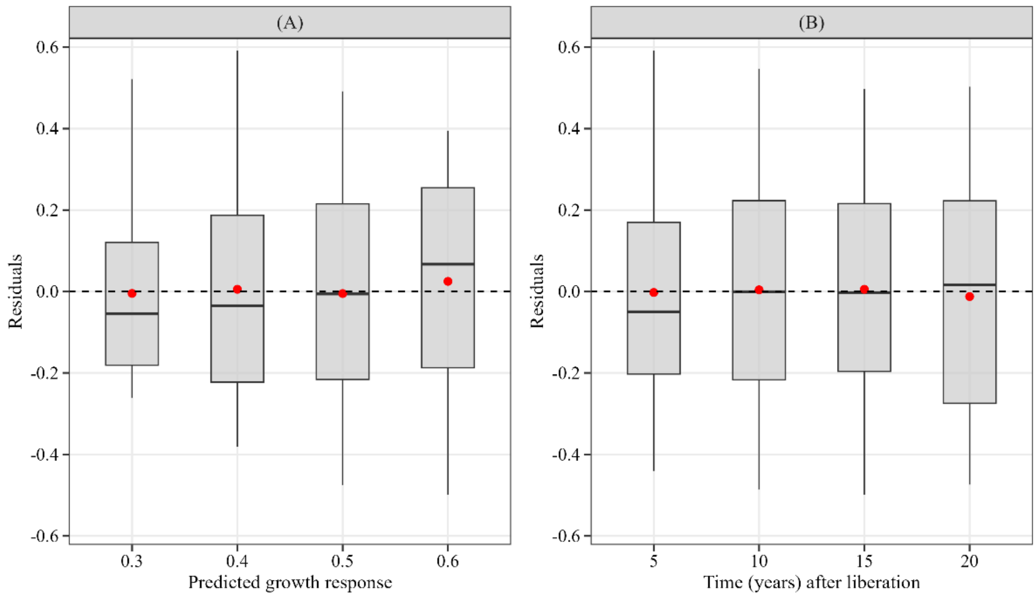


Figure 3. Residual analysis of the response model (Equation 1). Red dots are the means of the residuals.

Given the parameters of Equation 1 (Table 3), the expected growth response was simulated for a range of site indices (SIS) representing high and low site fertility to examine the suitability and reliability of the response model. As expected, the expected growth response after liberation was larger on a site with a low site index compared to a site with a high site index (Figure 4).

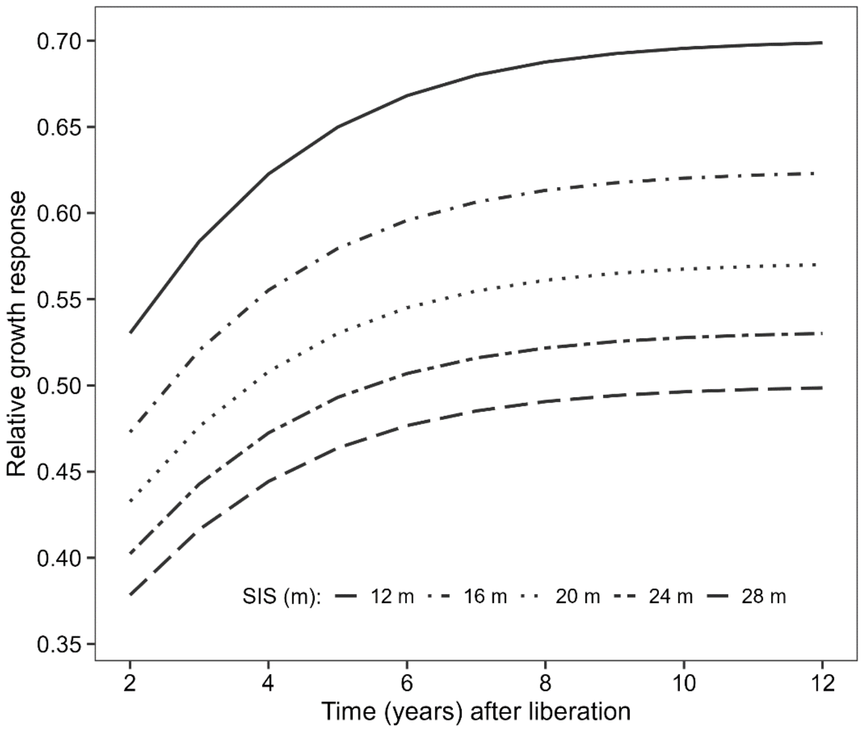


Figure 4. Simulated growth response for pines liberated on different site indices.

3.2 Basal area increment prediction

A two-step procedure was used to predict the annual increment of the retained pine tree's basal area (BAI). The first step concerned the development of the BAI model and in the second stage, the BAI model was integrated with the release response model developed in section 3.1. The parameters of the model predicting the annual BAI are given in Table 4. All parameters, except the intercept, were statistically significant at a 5% alpha level. The explained variance in BAI by the model (on the log-linear scale) was about 41% and the estimated residual standard error ($\widehat{\sigma}$) was 0.5045. The final predictors of the model included breast height diameter, crown ratio, site index, and qualitative information about stand management class, for example, unthinned stands and young stands above a height threshold of 3 m. The variance inflation factors (VIF) of the predictor variables were small suggesting low collinearity effects.

Table 4. Parameter estimates of the basal area increment model (Eq. 4). The dependent variable is ln(BAI). VIF is the variance inflation factor.

Variable*	Estimate	SE	T-value	p-value	VIF
Intercept	-0.2487	0.3592	-0.693	0.4890	
dbh^2	0.00042	0.0001	6.94	1.48e-11	2.116
CR	1.504	0.2362	6.367	5.62e-10	1.6498
SIS	0.0454	0.0084	5.427	1.03e-07	2.6129
$^{\mathrm{H}}/_{\mathrm{SIS}}$	0.6679	0.1905	3.505	0.000512	2.1874
$\mathrm{Y}_{\mathrm{ind}}$	-0.3240	0.6893	-4.700	3.65e-06	1.3049
$\mathrm{C}_{\mathrm{ind}}$	-0.5537	0.07523	-7.360	1.17e-12	1.3679
n	383				
ln(mean BAI)	2.6755				
$R_{adj.}{}^2$	0.4109				
$\widehat{\sigma}$	0.5045				

* See definitions of the parameters in the section 2.2.2. (n) is the number of observations used to fit the model.

Diagnostics of the model's residuals (Figure 5) showed homoscedastic patterns for predicted values of BAI and other variables such as latitude, soil moisture and the year of inventory. Figure 6 shows the predicted versus observed BAI in the original scale. Generally, the model shows moderate over-prediction for small and large true values of BAI (e.g., up to 30 cm²/year).

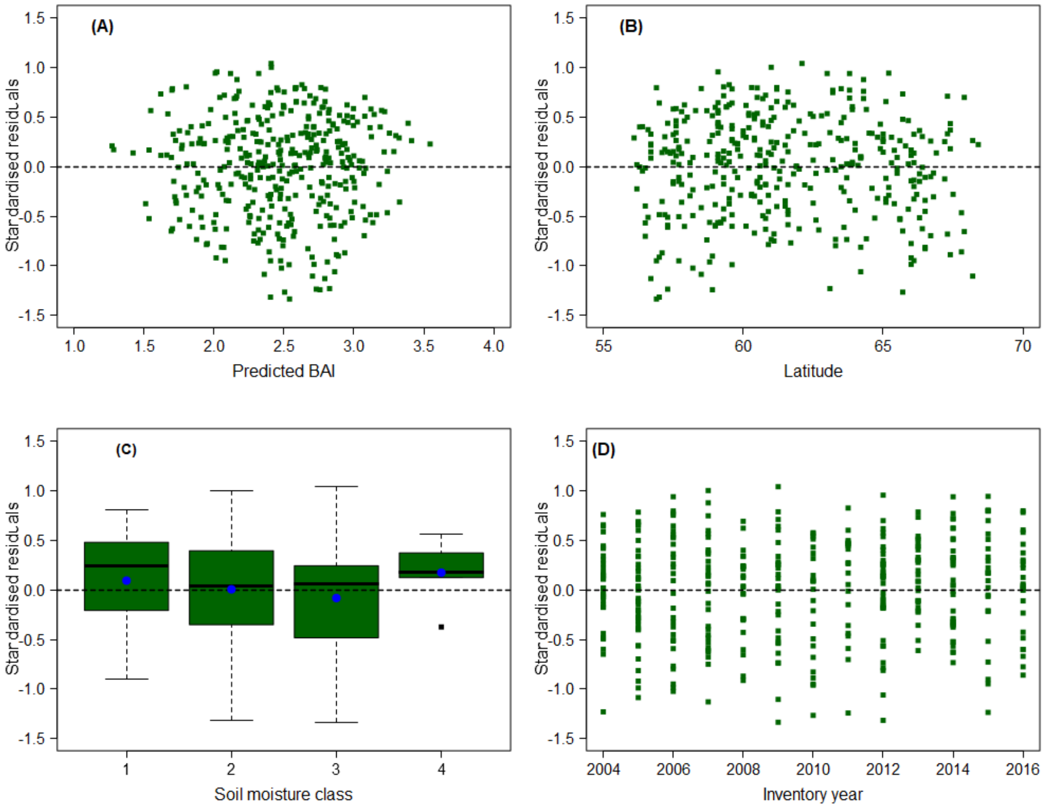


Figure 5. Diagnostic testing of the BAI model (Eq. 4). Residuals are presented in standardised form in the logarithmic scale. Blue dots in (c) represent the mean.

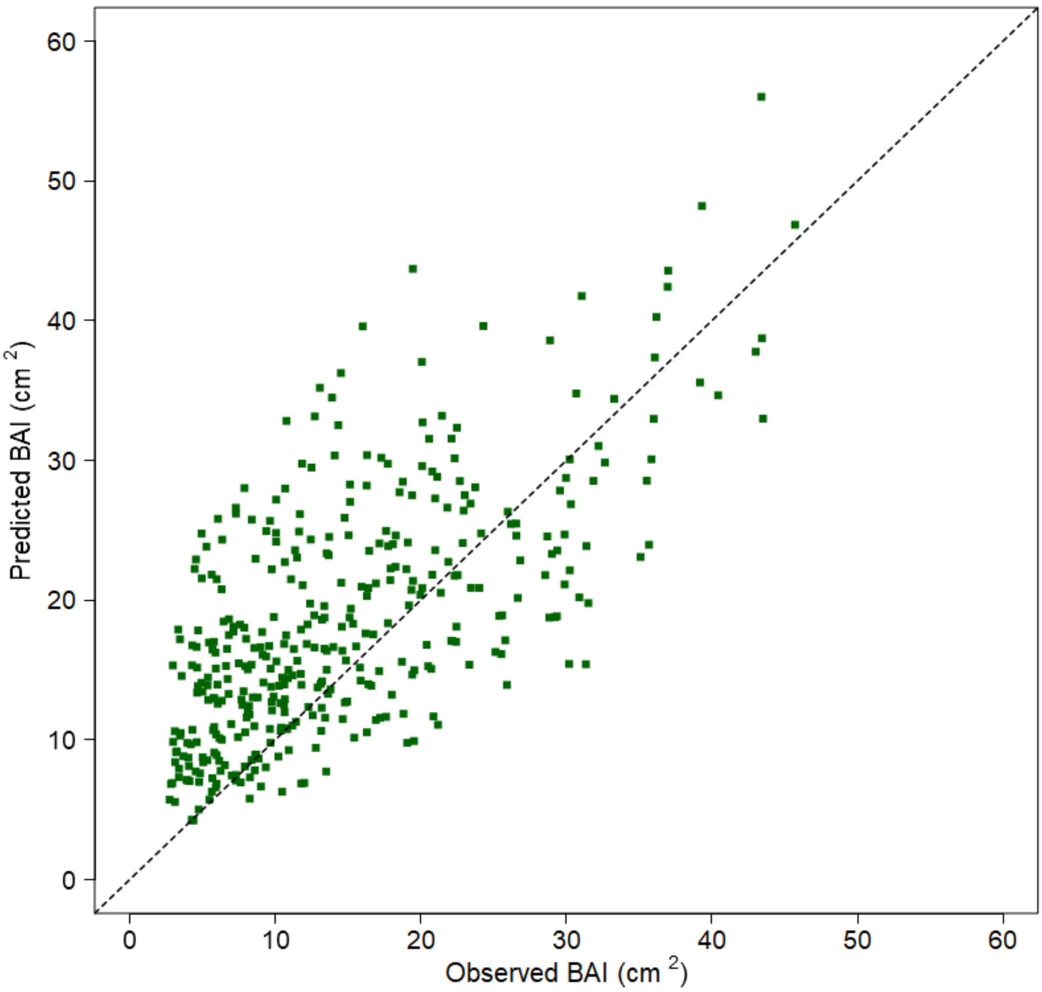


Figure 6. Observed versus predicted annual basal area increment (BAI) on the natural scale. The 1:1 line is shown as diagonal dashed lines.

In the second step, the release response model was incorporated into the general BAI model (Eq. 5). The final form of the integrated BAI model was structured as:

$$BAI_{adjusted} = \exp[f(\ )] * \exp\left(\frac{\hat{\sigma}^2}{2}\right) + \exp^{(\text{Liberation response})} \qquad (6.4)$$

In Eq. 6.4, the first term corresponds to the unadjusted growth according to Eq. 5. The last term defines the relative growth response after liberation (Eq. 1). To examine the feasibility of the integrated BAI model, predictions of BAI with and without liberation response were made using simulations. To begin with, intermediate models of dbh-age, height-diameter and crown ratio models were developed. These auxiliary functions were needed to update the values of the predictor variables in Eq.5 for every time point. The parameters and model behaviour of the functions are given in Appendix 1. The simulations were run for a high and low site index (28 m and 16m respectively) and at 100 years on annual time steps. To run the simulations, the following steps were implemented:

Step 1: For a given age (age at breast height) at liberation (i.e., age at release), compute the expected tree age for the other time points by adding the time after the liberation to the age at liberation.

Step 2: Compute the expected tree dbh given age (from step 1) and site index (determined by SIS and fixed for the whole simulation period) using the dbh-age model.

Step 3: From the tree dbh (given in step 2), compute the expected total tree height using the height-dbh model.

Step 4: From dbh, height and SIS, compute the expected crown-ratio using the crown-ratio model.

Step 5: Compute the growth reaction after liberation using Eq. 1.

Step 6: Predict the basal area increment (BAI) under the bark using Eqs. 4 and 5. This predicted growth is un-adjusted for the liberation response.

Step 7: Predict the adjusted growth using Eq. 6.4.

Given the adjusted (from step 7) and unadjusted growth (from step 6), the relative growth response was quantified as the ratio of the difference between adjusted and unadjusted growth to the unadjusted growth. The outcome of the simulations is shown in Figure 7. As expected, BAI increased with increasing diameter and the predicted BAI was lower independent on the site for the un-adjusted case (i.e., when the release response was excluded). The response reached its maximum within 5-10 years after liberation, and levels off by 50 years after liberation. The magnitude of the relative growth response was higher on poor sites than on rich sites.

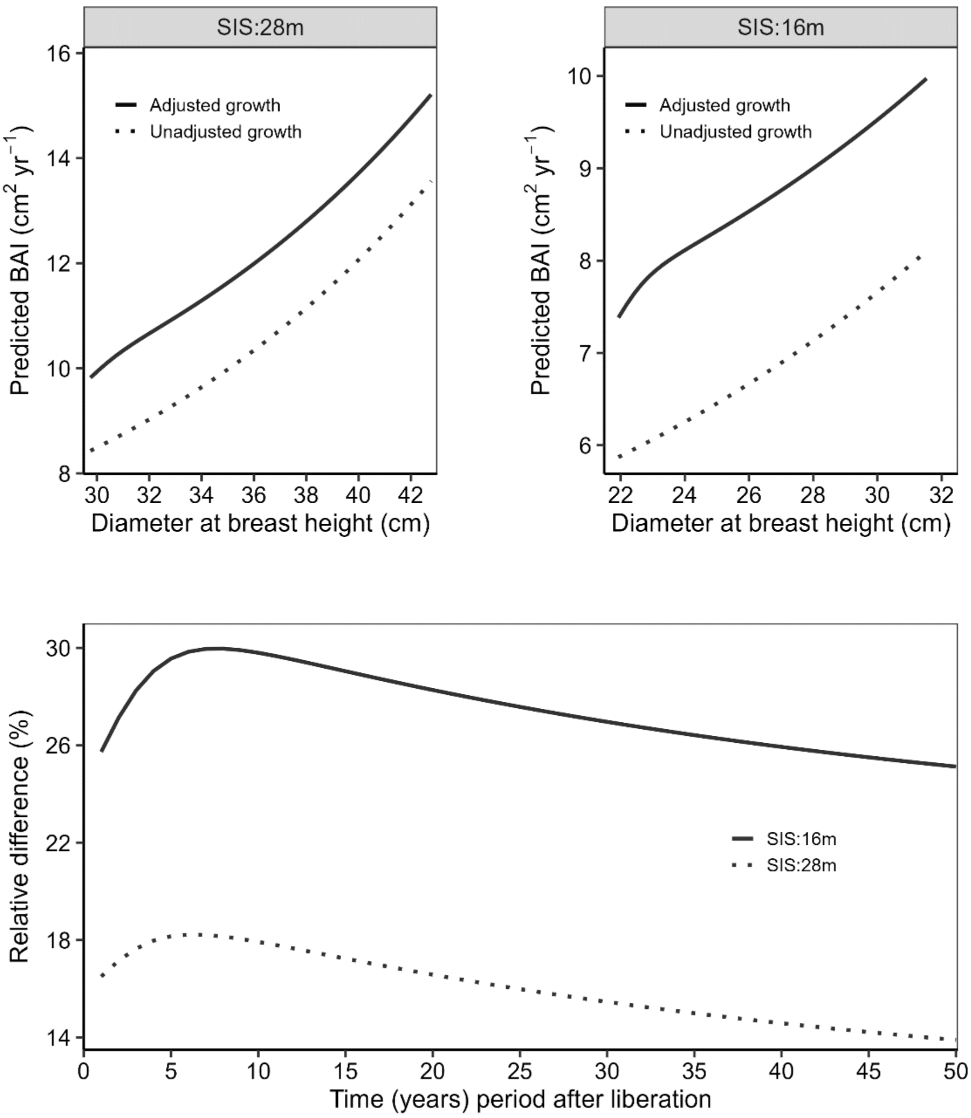


Figure 7. Simulations of annual basal increment of retained pine seed trees with and without adjusting for growth reaction after liberation on poor sites (SIS of 16 m) and rich sites (SIS of 28 m). Note the scale difference of the y-axes.

3.3 Tree damage prediction

The logistic model for predicting the probability of damage retained pine trees at time t was satisfactorily calibrated using the tree-level predictors of total height and age at breast height and site fertility variable described by site index (i.e., SIS). The collinearity among the predictors as measured by the VIF was under 2, suggesting negligible effects of multicollinearity. The estimated coefficients, standard errors and significance of the model parameters are given in Table 5 and the final form of the logistic damage model is presented:

$$\pi_{it} = \frac{e^{0.3838+0.0635*SIS-0.1299*Height+0.0091*Age}}{1+e^{0.3838+0.0635*SIS-0.1299*Height+0.0091*Age}} \qquad (7.5)$$

Table 5. Parameter estimates of the logistic model (Eq. 7.3).

Variable	Estimate	SE	Z-value	p-value	VIF
Intercept	0.3838	0.6439	0.596	0.5511	
SIS	0.0635	0.0310	2.048	0.0405	1.7855
Height	-0.1299	0.0322	-4.031	5.57e-05	1.8192
Age	0.0091	0.0029	3.086	0.0020	1.1527
n	330				

For inference of tree damage probability, the sign of the estimated coefficients was used to describe the direction of the relationship and the odds ratio to quantify the relationship. The marginal effects of the model predictor variables are simulated in Figure 8.

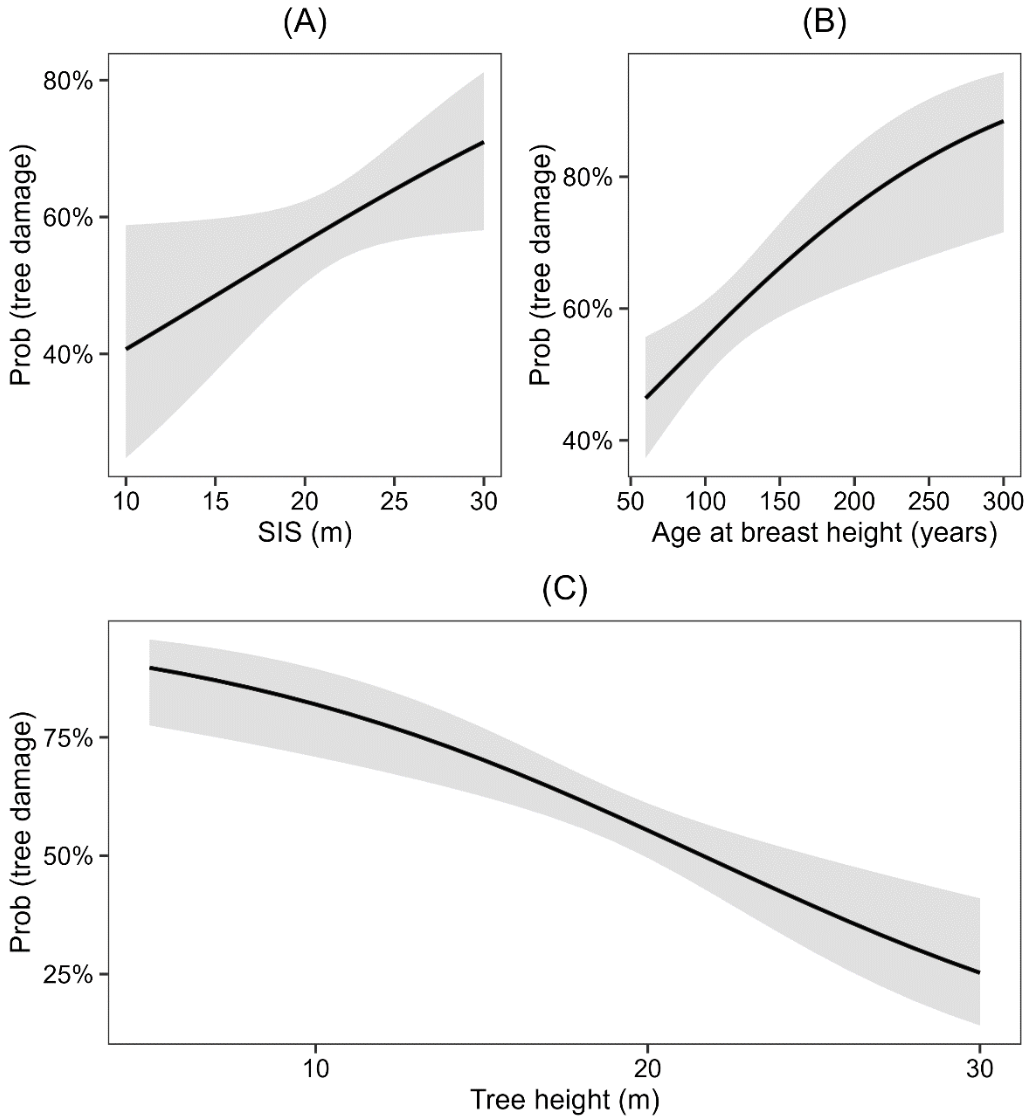


Figure 8. Simulated (A) marginal effects of site index (SIS), (B) tree age and (C) total height on the probability of a tree being damaged on measurement occasion t [i.e., $\Pr(y_{it} = 1)$].

Figure 9 shows the final predictors' estimated odds ratio and 95% confidence intervals. The site index and age parameters were positive, suggesting a positive association with the probability of a tree being damaged on measurement occasion t. When the site index increased by 1 m, the probability of a tree being damaged increased by 6.56%. Similarly, an extra year of age increased the odds of a retained pine tree damage probability by a factor of 1.0091 units (thus, representing an increase in the damage probability by 0.91%). However, the estimated coefficient of total tree height was negative, suggesting decreased odds of a tree being damaged. The chance of a retained pine tree being damaged decreased by 12.18% with a 1 m increase in height.

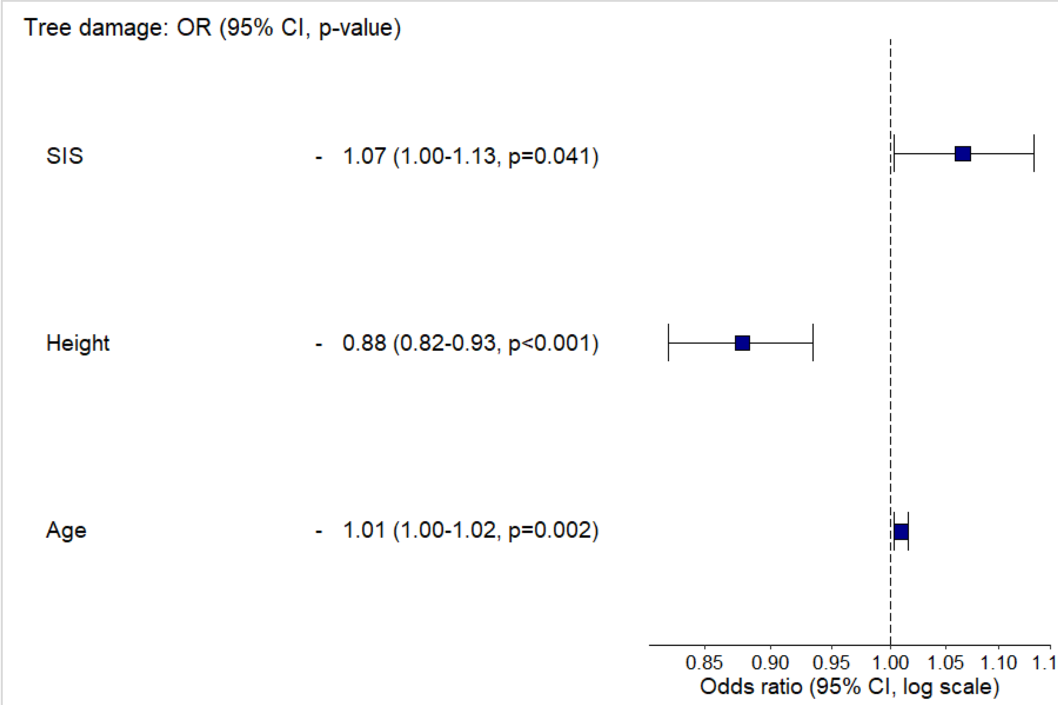


Figure 9. Odds ratio (OR) and 95% confidence intervals describing the effect of site index (SIS), age (number of rings at breast height) and tree height on damage probability of retained pine trees.

From the receiver operating characteristics curve, the accuracy of the model determined as the area under the curve (AUC) was about 66% (Figure 10). By translating the probability of damage predictions into discrete events using a cutpoint value of 0.5, the sensitivity and specificity measures of the model were quantified (see the confusion matrix in Appendix 2 for the observed and predicted tree damage using a cutpoint value of 0.5). The sensitivity of the model was 80% indicating that the model correctly predicted damaged trees as being damaged. Nevertheless, the model's specificity was low, as only 40% of the undamaged trees on measurement occasion t [i.e., $\Pr(y_{it} = 0)$] was correctly identified as not damaged by the model.

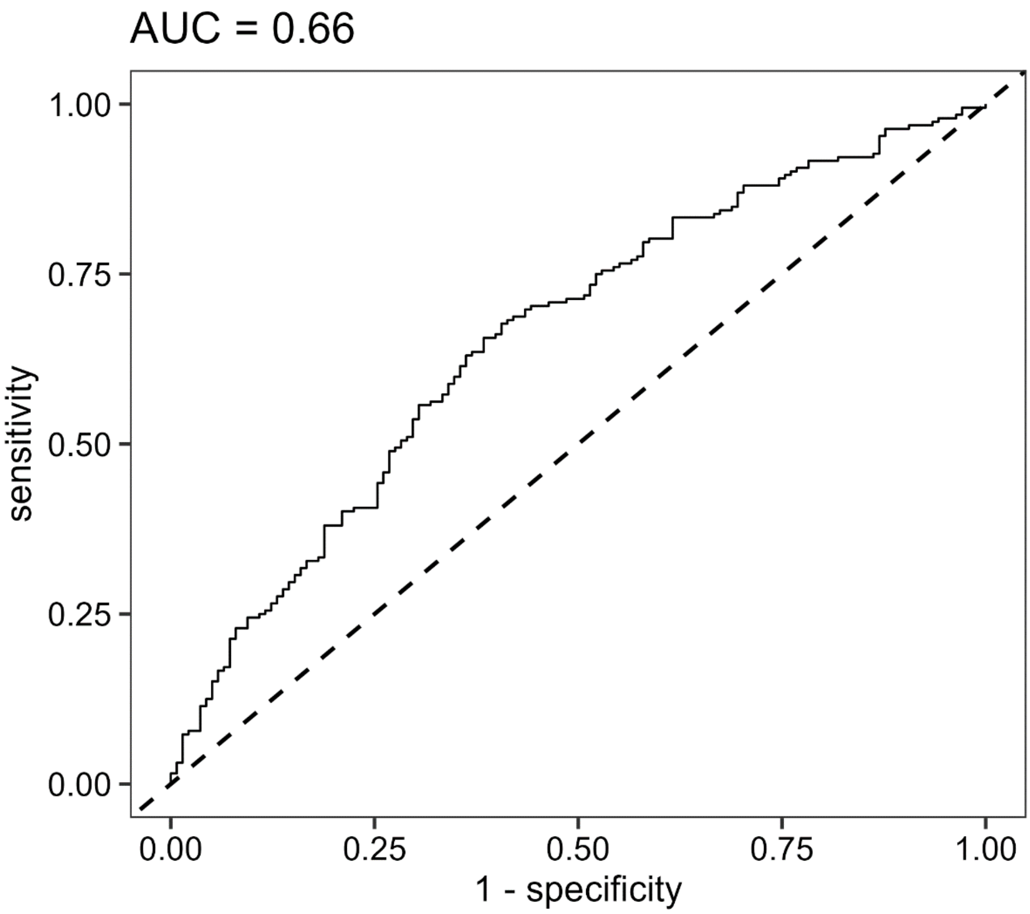


Figure 10. Accuracy of the logistic model predicting the probability of retained pine tree damage on measurement occasion t [i.e., $\Pr(y_{it} = 1)$]. Accuracy is measured by the area under the receiver operating characteristic curve (AUC).

4. Discussion

In this study, the growth of retained pine seed trees has been investigated. In particular, aspects of liberation response and annual basal area increment underbark were examined using statistical modelling. Other aspects concerning damage risks were also explored. Data from the temporary plots of the Swedish NFI facilitated the investigation.

Tree diameter growth is known to be affected by stand density. Thinning decreases stand density abruptly, which strongly impacts tree growth (Hynynen, 1995). Growth response following thinning is a result of (i) increased growing space, (ii) the fertilisation effect provided by the non-harvested parts of felled trees, and (iii) the selection effect where trees retained in stands subjected to thinning from below have grown better before thinning compared to those removed during thinning (Hägglund, 1981). These effects are widely known in young and established stands (Assmann, 1970). Similarly, following a clear-cut, retained trees also react to available space and nutrients needed for growth.

Jonsson (1995) defined the thinning response of a tree as the ratio between the actual annual ring width and the corresponding assumed annual ring width unaffected by thinning. This work defined the reaction response after thinning (i.e., liberation) differently. The corresponding unliberated trees (i.e., control) were unavailable for a given retained pine tree. Instead, the width of the annual rings several years before and after liberation was used to examine the effect of liberation on the diameter growth of retained pine trees. Subsequently, the growth response was defined as the basal area increment after liberation relative to growth at the time of liberation (i.e. the year of stand harvest). The retained pines had been liberated at different times. Determining the harvest year from survey data (such as those from NFIs) is difficult and may introduce large uncertainties in quantifying the growth response of liberated trees. To overcome this problem, auxiliary information such as the age of the new stand and patterns of the annual ring widths before and after liberation was used. This approach was successfully used by Jakobsson (2004) to investigate the growth of retained pine trees in Sweden.

The pattern of the liberation response (Figs. 2 and 4) was similar to earlier reports for pine and spruce (Jonsson, 1995; Jakobsson, 2004). The magnitude of the relative growth increase 10 years before and after liberation was 11% and 21%, respectively, for the high and low site index (Table 2). The duration also lasted for more than 20 years after liberation. The logistic function satisfactorily predicted the liberation response in annual time steps and only required the site index as the predictor variable. The release response is bounded between 0 and 1 indicating a direct proportional effect of the liberation response.

In many growth simulators used in forest management planning, the effect of thinning is accounted for by incorporating an explicit thinning variable in the growth model (e.g. Hynynen, 1995). In this study, the liberation response model was included as an additive component (Eq. 6.4) to a general basal area increment prediction model (i.e. adjusted model). In this way, it is possible to isolate the liberation response from the reference growth model that generally depicts tree basal area growth in stands before harvest. The adjusted model predicts the growth in annual time steps even though age is not required in the model. Age and size (e.g., diameter) are highly correlated and as such, the use of diameter is preferred due to its availability and less error-prone compared to age. Jonsson (1995) found very little effect of tree age in the thinning response functions of pine in Sweden. Näslund (1942) also reported that tree age does not have any clear (According to model (6.4) and Figure 4, liberation increases the relative basal area growth without delay. The response reaches its maximum within 5-10 years after thinning, and levels off by 30 years after liberation (Fig. 7). Similar results were obtained by Hynynen (1995) in predicting the growth response to thinning for pine in Finland. To facilitate the implementation of the adjusted growth model, supplementary functions were also derived. These involved models for diameter and height development, and crown ratio (see Appendix 2). These auxiliary functions are useful for updating the values of the predictor variables in Eq.5 for every time step.

In addition, damage prediction models were developed. However, the model has moderate accuracy, especially in predicting the status of actual undamaged trees. One of the main reasons is the coarse categories of damage used in the analyses. The NFI records about 11 damage classes; and these were further reclassified into two groups: damaged and undamaged trees. These individual damages likely have different underlying factors and as such, the predictors used in the model may fail to capture. The most prevalent damage for retained pine trees is windthrow. Tree dimensions such as height, diameter and crown volume are strong predictors (Valinger and Fridman, 1999). In the damage models, the final predictors included height, site index and age (Eq. 7.5). Others such as crown ratio and diameter were statistically not significant when tested. In addition, the model suggested that the chance of a retained pine tree being damaged decreased by 12% with a 1m increase in height. Two potential explanations for this are (i) the culmination of height growth in older trees and (ii) the deeper root system of pines making them less susceptible to windthrow (Karlsson et al., 2017).

5. Conclusion

This project investigated the basal area growth and damage of retained pine seed trees in Sweden. About 400 sample plot data from the temporary plots of the Swedish NFI facilitated the investigation. A two-step procedure was used to model the annual basal area increment under bark. First, the release (i.e., liberation) response of the annual ring-width increment in single trees of retained pine seed trees was examined and a logistic regression function was developed further to quantify the response. The results of the response examination show that there is a significant increase in the annual basal area growth, which is persistent up to 20 years after liberation. The magnitude of the response depended on the site index. On sites with high site index, the relative increase in growth 10 years before and after liberation was 11%. The corresponding value for sites with low site index was 21%. The maximum relative response culminated at six and seven years after liberation for high and low site indices, respectively.

In the second stage, the liberation response model was integrated into a model predicting the annual basal area growth. The new growth functions adapted well to calibration data. Likewise, the model seemed flexible to apply and simulated outcomes demonstrate logical predictions.

In tandem, damage models were also explored. A logistic regression function was used to predict the probability of retained pine trees being damaged at a given time given tree height, age, and site index information. The model accuracy using a cut-off value of 0.5 was moderate (i.e., 66%), correctly predicted damaged trees (sensitivity of 80%) but poorly predicted undamaged trees (specificity of 40%).

In conclusion, the models are ready to be implemented in the Heureka Decision Support System for simulating retained pine tree annual growth and damage. It is recommended that a statistical sensitivity analysis of the functions be conducted to assess the impact of the renewed models by evaluating predictions of mean basal area growth (and volume growth) and their confidence intervals for large areas.

References

Assmann, E. (1970). The principles of forest yield study. Pergamon Press, Oxford, New York, pp. 506.

Baskerville, G. L. (1972). Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research*, *2*(1), 49-53.

Elfving, B. (2010). Function for time to reach breast height. SLU, PM for Heureka.

Fridman, J., Holm, S., Nilsson, M., Nilsson, P., Ringvall, A.H., Ståhl, G. (2014). Adapting National Forest Inventories to changing requirements - the case of the Swedish National Forest Inventory at the turn of the 20th century. Silva. Fenn. 48, 1095. https://doi.org/10.14214/sf.1095.

Hagberg, N. (1942). Stamformens förändringar hos tall och gran under beståndsutvecklingen och efter friställning. Svenska skogsvårdsföreningens tidskrift (1), 44. In Swedish with German summary.

Hägglund, B. (1981). Forecasting growth and yield in established forests. An outline and analysis of the outcome of a subprogram within the HUGIN project. Swedish University of Agricultural Sciences, Department of Forest Survey, Report 31. 145 p.

Hägglund, B., Lundmark, J.E. (1977). Site index estimation using site properties. Scots pine and Norway spruce in Sweden, Studia Forestalia Suecica, p. 38.

Hynynen, J. (1995). Predicting the growth response to thinning for Scots pine stands using individual tree growth models. Silva Fennica 29(3): 225-246.

Jakobsson, R. (2005). Growth of retained Scots pines and their influence on the new stand. Doctoral thesis. SLU, Umeå.

Jonsson, B. (1995). Thinning response functions for single trees of Pinus sylvestris L. and Picea abies (L.) karst. Scand. J. For. Res. 10, 353–369. https://doi.org/10.1080/02827589509382902

Karlsson, C., Sikström, U., Örlander, G., Hannerz, M., Hånell, B., Fries, C. (2017). Skogsskötselserien nr 4, Naturlig föryngring av tall och gran. Skogsstyrelsen. Jönköping, Sweden. pp. 1-144.

Karlsson, C., & Örlander, G. (2004). *Naturlig föryngring av tall*. Skogsstyrelsen. Jönköping, Sweden. pp. 1-90.

LeBlanc, D.C. (1990). Relationships between breast height and whole-stem growth indices for red spruce on Whiteface Mountain, New York. Can. J. For. Res. 20, 1399–1407. https://doi.org/10.1139/x90-185.

Mensah, A. A., Petersson, H., Dahlgren, J., Elfving, B. (2023). Taller and slenderer trees in Swedish forests according to data from the National Forest Inventory. Forest Ecology and Management, 527, 120605.

Näslund, M. (1942). Die reaktionsfähigkeit des alten norrländischen fichtenwaldes nach durchhauung. Meddelanden från Staten Skogsförsöksanstalt, 33(1).

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www. R-project.org/.

Söderberg, U. (1992). Funktioner för skogsindelning. Höjd, formhöjd och barktjocklek för enskilda träd. SLU, Rapporter – Skog. Inst. För skogstaxering, Rapport 52.

Schultz, R. P. (1997). *Loblolly pine: the ecology and culture of the loblolly pine (Pinus taeda L.)* (No. 713). US Government printing office.

Valinger, E., Fridman, J. (1999). Models to assess the risk of snow and wind damage in pine, spruce, and birch forests in Sweden. *Environmental Management*, *24*, 209-217.

Valinger, E. (1992). Effects of wind sway on stem form and crown development of Scots pine. Australian Forestry 55, 15-21.

Zeide, B. (1993). Analysis of growth equations. Forest science, 39(3), 594-616.

Acknowledgements

This report was written by Alex Appiah Mensah on behalf of the the Swedish Environmental Agency which financed the work (Naturvårdsverket, NV-06834-23). The author is responsible for the content, conclusions and recommendations. Lars Sängstuvall initiated the project and Jonas Dahlgren from the Swedish National Forest Inventory (Department of Forest Resource Management, SLU) provided the data. Jeannette Eggers was the project administrator and commented on the initial draft of the manuscript.

Appendix 1: Auxiliary functions

Diameter-age model:

The function is used to predict tree diameter at breast height (dbh) over bark using age at breast height and site index (SIS). The generalised Chapman-Richards model was used as the functional form and the parameters were estimated via generalised nonlinear least squares. The parameters were all significant at the alpha value of 0.05. The model has a residual standard error of 7.118 cm and the form is given as:

$$dbh = (0.5449 * SIS^{7.1906}) \cdot (1 - e^{-0.0235*Age})^{1.1072}$$

Figure A1 shows the model and data.

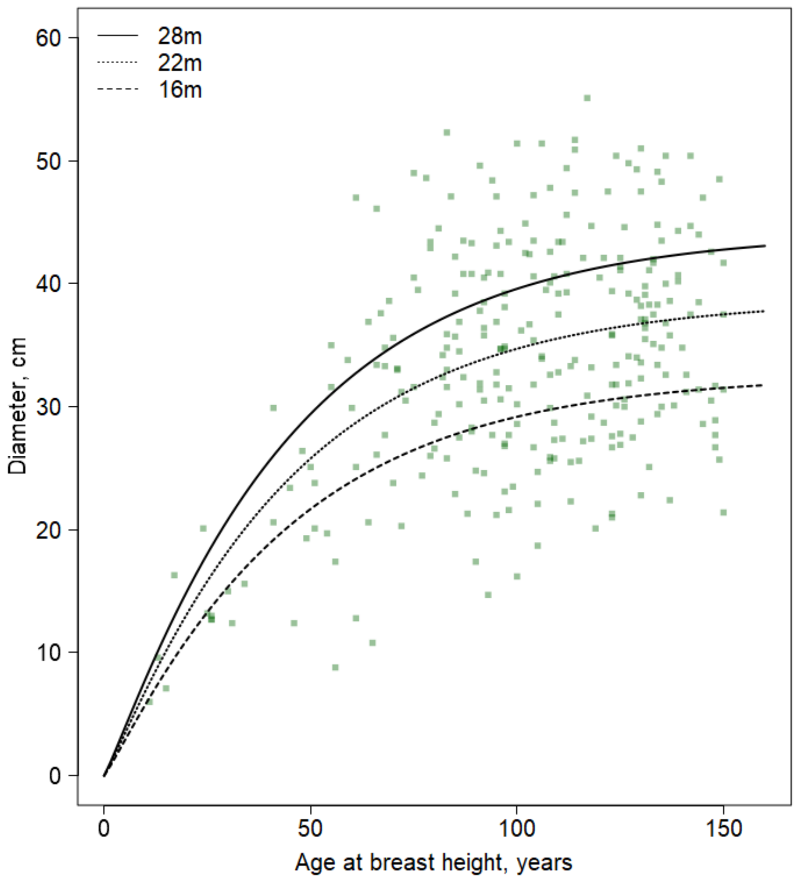


Figure A1. Dbh-age relationship for retained pine trees. The data is shown as green squares and model-based predictions are shown with lines for different SIS values.

Height-diameter model:

The Näslund's function was used to model the relationship between the height and diameter of pine seed trees. The parameters of the model were estimated using non-linear least squares. The model had a residual standard error of 3.37 m and the form was:

$$h = 1.3 + \frac{dbh^2}{(2.4671 + 0.1598 * dbh)^2}$$

The data and model are shown below.

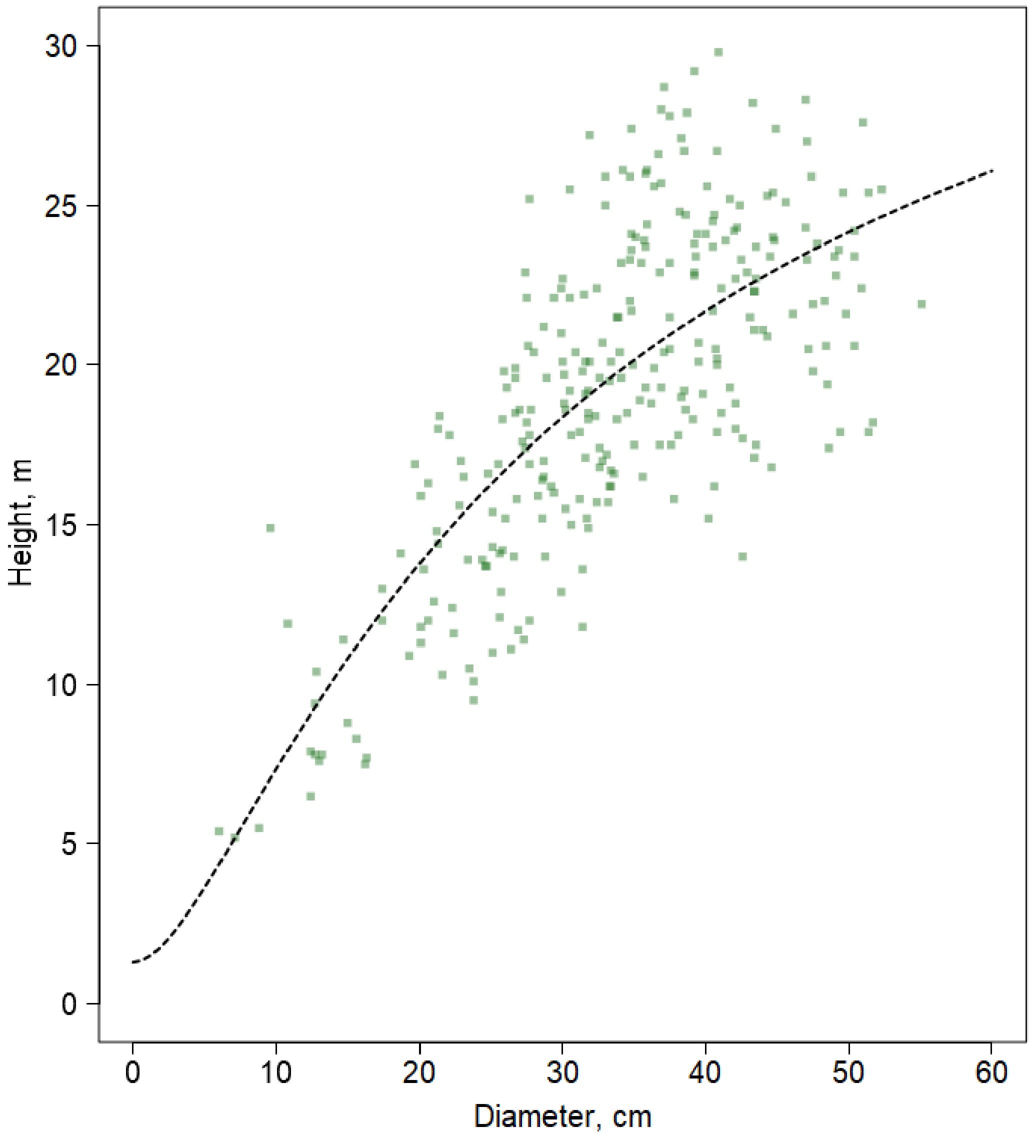


Figure A2. Height-diameter relationship of retained pine seed trees. The data is shown as green squares and the model is given as dashed lines.

Crown ratio model:

A logistic function was used to relate the mean crown ratio (CR, expressed as the ratio of crown length to total height) to dbh, height (h) and site index (SIS). The parameters were estimated by nonlinear regression and were all statistically significant at a probability of 0.05. The model had a residual standard error of 0.1122. The explicit form of the model was:

$$CR = \frac{1}{1 + e^{-(1.846+0.00029*dbh-0.0805*h-0.02502*SIS)}}$$

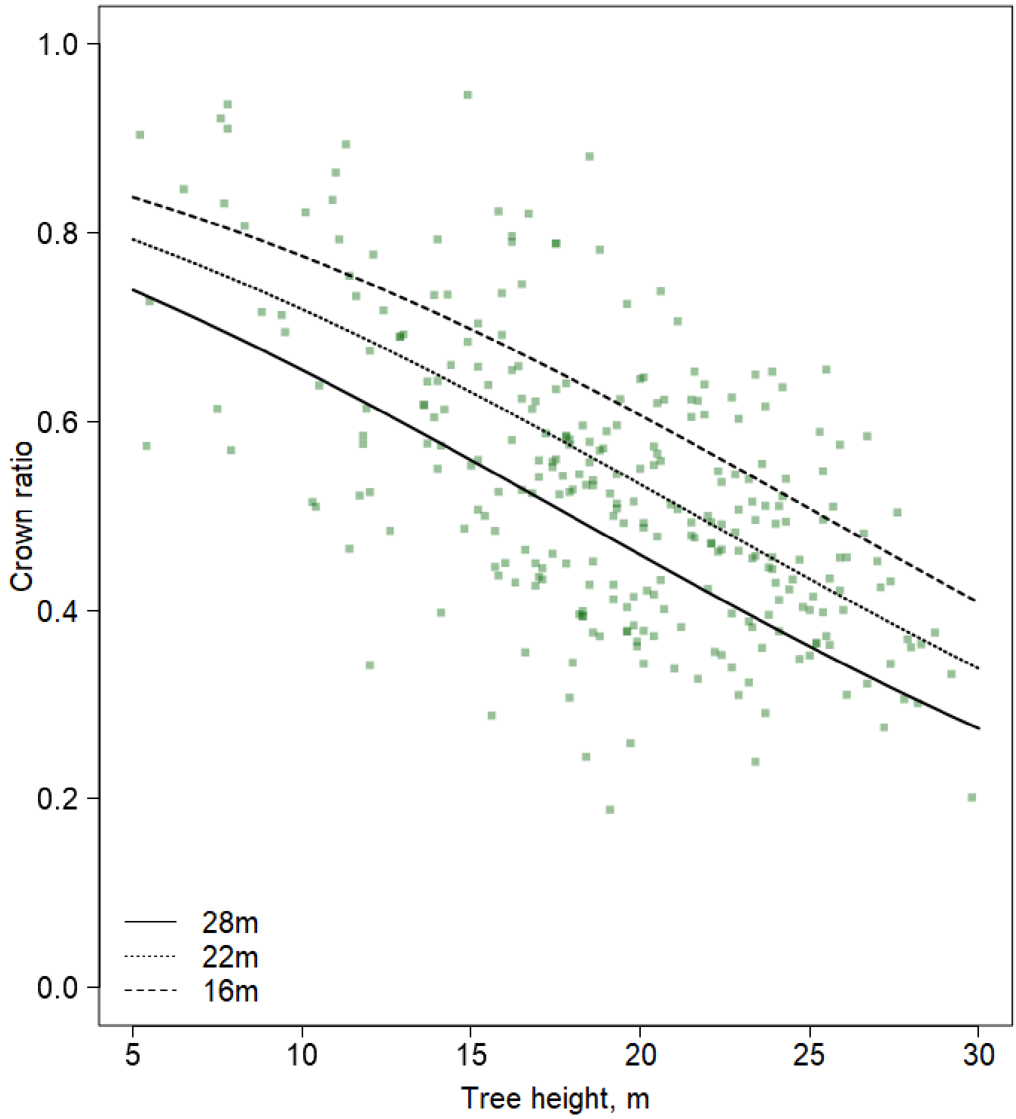


Figure A3. Model-based predictions of crown ratio retained pine seed trees given height, diameter and site index. The diameter was set to 30 cm (at age 150) for the SIS 16m, 37cm for (at age 140) SIS 22m and 42cm (at age 100) for the SIS 28m.

Appendix 2: Accuracy of damage models

The section describes the accuracy of the model for predicting damage of retained pine seed trees at a given time t. The general model is given in Equation 7.5 and the accuracy is shown in Figure 10. Using a cutpoint of 0.5, the probabilities were discretized, i.e. damage or no damage. The confusion matrix of the observed and predicted damage outcomes is presented below:

Table A2: Confusion matrix table for damage predictions of pine trees.

	Predicted	
Observed	No damage	Damage
No damage	57	81
Damage	38	154

From Table A2, the sensitivity and specificity metrics are derived. Sensitivity measures the ability of the model (given a threshold value of 0.5) to detect the condition when the condition is present. In this case, the percentage of damaged trees is correctly identified as being damaged. The sensitivity of the model was 80 %.

$$\text{Sensitivity} = 154/(154 + 38) = 0.802$$

On the other hand, specificity deals with quantifying the model's ability to correctly exclude the condition when the condition is absent. In this case, it is the percentage of undamaged trees which are correctly identified as not damaged. The specificity of the model was 41 %.

$$\text{Specificity} = 57/(57 + 81) = 0.413$$