Basal Area Growth and Yield Models for Siberian Larch

Alfred Deutgen

Master's thesis • 30 credits
Swedish University of Agricultural Sciences, SLU
Southern Swedish Forest Research Centre
Jägmästarprogrammet
Alnarp 2023

Basal area Growth and Yield models for Siberian Larch
Grundytetillväxt- och utvecklingsmodeller för Sibirisk lärk

Alfred Deutgen

Supervisor: **Urban Nilsson, Swedish University of Agricultural Sciences, Southern Swedish Forest Research Centre**

Examiner: Per-Magnus Ekö, Swedish University of Agricultural Sciences, Southern Swedish Forest Research Centre

Credits:	30 credits
Level:	Second cycle, A2E
Course title:	Masterarbete i skogsvetenskap
Course code:	EX0985
Programme/education:	Jägmästarprogrammet
Course coordinating dept:	Southern Swedish Forest Research Centre
Place of publication:	Alnarp
Year of publication:	2023
Cover picture:	Alfred Deutgen
Copyright:	All featured images are used with permission from the copyright owner

Keywords: Growth model, Yield model, Siberian Larch, Russian larch, *Larix sibirica*, *Larix sukaczewii*, Lundqvist-korf model, Mixed effect model

Swedish University of Agricultural Sciences
Faculty of Forest Sciences
Southern Swedish Forest Research Centre

Abstract

There is no species-specific basal area growth model for Siberian Larch in Sweden, which is necessary in evaluating forest management decisions. Three different models were created from stand trials around Sweden. An individual tree basal area growth model (BAI1), a per hectare basal area growth model (BAI2) and a model for total basal area production (BAY). The individual tree basal area growth and the total basal area production models were converted to models for basal area growth per hectare, so that all three could be compared. The BAI2 model gave the best result of the models and showed a clear improvement compared to the previous model used for Siberian larch.

Keywords: Growth model, Yield model, Siberian Larch, Russian larch, *Larix sibirica*, *Larix sukaczewii*, Lundqvist-korf model, Mixed effect model

Table of contents

List of tables

List of figures

Introduction

1.1 Background

Siberian larch is the only larch species that is native to Sweden and together with Scots pine and Norway spruce, are the only three native coniferous species suitable for production forestry in Sweden. The Swedish forestry legislation is the same for Siberian larch as the other native planted conifers (Normark & Fries 2019). The use of non-native tree species must be reported to the Forest Agency and are restricted close to nature reserves (Norén 2009). In addition, if the estate is certified according to FSC or PEFC, more than a specified percentage of the forest land cannot be planted with non-native tree species. Together with all the other larch species available for planting, Siberian larch only makes up 0.1% of the total national forest volume (Nilsson et al. 2021). There is potential of up to 2.9 million hectares of suitable land to plant Siberian larch in northern Sweden (Söderholm & Öhman 2010). However, there is little information how the stands develop over time. Estimation and prediction of future stand properties is important for effective management of forests. Models for growth and yield are necessary for choosing tree species for planting, evaluating which forest management actions should be taken and when interventions should be done (Heiðarsson & Pukkala 2012). Currently, there is no basal area growth function specified for Siberian larch in Sweden (Carlzon 2022). In the decision support system Heureka, the model that has been utilized for estimating basal area growth for Siberian larch, is a model created for Scots pine (Carlzon 2022). Carlzon (2022) hypothesize that the estimations for Siberian larch could be improved by creating a species-specific growth function. The Scots pine model have been shown to over- and underestimate basal area growth for Siberian larch, which may result in high uncertainty in predicting stand-level growth. Having an accurate basal area growth model is essential in planning thinning regimes and being able to compare management options and systems.

1.2 Siberian Larch

There is evidence that native Siberian larch grew in the Swedish mountains 7500-8700 years ago, but has not been continuously present since then (Kullman 1998; Karlman 2010). Imported Siberian larch has been planted and managed in Sweden for more than 200 years (Martinsson 1995). They usually replace a pine or spruce forest and is managed using a classic clear-cut silvicultural system. In the 1930s and 1950s imported plant material from two areas in central Russia had a high mortality, and those that survived developed low quality, which reduced the interest in the Siberian larch (Karlman 2010).

The Siberian larch forests in Sweden can be assumed to be either Siberian larch (Larix sibirica) or Russian larch (Larix sukaczewii). Siberian- and Russian larch are similar enough that there is no consensus on whether they are two different species or not (Martinsson 1995). In Sweden they are mostly seen as the same species, but other Nordic countries have made the decision to count them as two different species (Pesonen et al. 2009). Larch that has come to be planted in many places in the Nordic countries has been called Siberian larch even though it comes from the Russian larch regions (Karlman 2010). There are some differences between the two larch species, L. sukaczewii originates from western Russia to the Ob River, in western Siberia. Further east, L. sibirica becomes the dominant tree species of the two. For Scandinavian conditions, L. sukaczewii has been shown to have lower mortality and higher growth than L. sibirica. Larix sibirica thrives in continental climates but is less adapted to maritime climates where its more susceptible to damage caused spring frosts (Abaimov et al. 1998). Provenance selection is however a probable reason for this, as L. sukaczewii comes from a climate more similar to that of Sweden. It is important for all tree species that the right provenance is chosen for the planting site to assure a high survival and growth. In this thesis, the two larch species are not distinguished from each other and is henceforth called Siberian larch.

Siberian larch is a fast growing pioneer species (Arvidsson 2006). The most common management practise is a classic clear-cut silviculture system with 4-5 thinnings and an early clearcuts at about 55 years of age, if the stand is planted on medium fertile soil. If high dimensional timber is the goal one can wait 20 years more.

Siberian larch has a growth rate that can compete with Norway spruce but differs in the influence of biotic and abiotic factors compared to both Norway spruce and Scots pine (Martinsson 1995). This makes Siberian larch an attractive addition to the forests of northern Sweden given the uncertainty about climate change. Larch can withstand higher browsing pressure than pine and is generally more storm-resistant than spruce (Westin et al. 2016). It has been hypothesized that deciduous trees are less prone to storm damage since they are leafless during the autumn and winter, which is when most heavy storm damage occurs in Sweden (Valinger & Fridman 2011). The loss of needles in winter suggests that it may also positively affect the soil for tree growth (Karlman 2010). The soil can quickly develop a thicker snow cover in winter, preventing it from freezing, and in the spring, the soil will also thaw and warm up faster. This can also result in differences in ground vegetation compared to pine and spruce on similar soil, which could contribute to higher diversity. Sometimes a thick mat of needles is formed on the ground under larch stands which could reduce or hinder growth of ground vegetation. A growth-model that includes ground vegetation as a variable, without including the fact that they receive more sunlight due to the larch's nature, could systematically overestimate or underestimate growth. A sensitivity analysis showed that using vegetation in the model predicted with more error than not using vegetation (Carlzon 2022).

1.3 Growth and Yield models

The task of Growth and Yield studies is to examine a process in the forest, in relation to time and some form of technical measurement, like volume or basal area (Assmann 1970). By measuring a tree before and after a certain period of time, it is possible to calculate its incremental growth. Yield is usually seen as the total harvested volume plus volume available for harvest at a given time, in this case however it will be basal area. There are many factors influencing tree growth and yield, which makes the models created not based on a strict biological relationship, but a stochastic relationship based on probability. Growth models can be derived from yield models, and vice versa (Goude et al. 2022).

1.3.1 Regression models

It is common to use linear-regression to create growth models (Goude et al. 2022). They are usually created with inventory data and has proven to be accurate in their estimations. They are robust and can be made with easy to measure variables, so the model itself can be easily utilized (Weiskittel et al. 2011). With a large dataset covering different managed forests, there is a possibility to include management interventions in the model. However, variables not represented in the data cannot be accounted for. If the data comes from a stand that has been managed with a thinning regime, predicting natural mortality will be difficult, as the trees that would have died were taken out before self-thinning.

1.3.2 “Lundqvist-Korf” function

The Yield model in this study is made by using a non-linear asymmetrical sigmoid function, a S-curve. The function that have been used in this study is the “Lundqvist-Korf” function (Korf 1939; Lundqvist 1957; Panik 2013). The relationship of the variables is already in an equation, but has two variables that needs estimation. The yield model has the aim to predict the total basal area production, given a starting year and total basal area. This includes all basal area removed in thinnings. It is a function that is relatively simple that can be used for different types of yield models, such as height, volume, and basal area (Feng-ri et al. 2000; Sedmák & Scheer 2015). The function was used for a basal area yield function for birch in Sweden (Liziniewicz et al. 2022). Being a relatively simple function that has been shown to work in Sweden for both deciduous and coniferous trees made it a good candidate for this study.

1.4 Aim of thesis

At this moment there are no species-specific growth and yield model for basal area in Siberian larch in Sweden. The models used for estimating today are the same as for Scots pine (Carlzon 2022). Since the interest in Siberian larch has increased in northern Sweden, it is important that simulations of future growth are as accurate as possible. Therefore, the aim of this thesis was to evaluate possible improvement of growth models used in the Heureka decision support system. Three models for Siberian larch, basal area growth for individual trees, per hectare basal area growth and total basal area production per hectare was created and compared to each other and to current models used in the Heureka system.

2. Materials and method

Three types of Siberian larch models were created. Basal area increment for individual trees (BAI1) and per hectare (BAI2), as well as basal area yield for total basal area production per hectare (BAY). This was done using data from long-term experiments in Siberian larch located at five sites in Sweden (Figure 1). In each site, several plots with different treatments were measured several times. Details regarding each site is found in Table 1. The plots have been established as forest trials, managed with a classical clearcut silviculture system. The origins of the plant material is mostly unknown, making it possible to be from either L. sibirica or L. sukaczewii.

Table 1. Information regarding about the dataset and each site. Site index is not preset here as it needs to be estimated, see 2.1.1.

Site	Plots	Number of plot revisions	Individual trees measured	Number of Indidiual tree revisions
5	26	74	1662	6772

Site	*Kvisseln*	*Vindeln 1*	*Vindeln 2*	*Dorotea*	*Siljanfors*
latitude	58,3	64,3	64,7	64,7	60,9
longitude	14,7	19,8	19,3	15,3	14,3
Plots	1	3	3	6	12
Area of plots, m2	1500	1156	1490	800-1000	225-2000
Altitude (m,a,s,l)	175	300	315	435	220-370
Number Revisions	2	2	1	2	1-5
Age range	73-97	19-29	17-27	33-55	19-96

Figure 1. Map over Sweden where the sites are located.

2.1 Basal area increment model methodology

Linear mixed-effect model parameters were estimated using the program R (4.2.2) with R Studio (Posit team 2023) with lme4-package (Bates et al. 2015). Choosing the right random effect for BAI1 and BAI2 with different nesting levels must be considered. The individual trees and plots have been measured between 2-6 times. For BAI1 the random effect should therefore be at individual tree level. The BAI2 model utilizes several plots from different sites. The plots in the sites have some differences, like altitude, making it adequate to set the random effect on plot level. The random effect helps to avoid problems of differences in the amount of data from the various sites, creating false positives in the model, like high p-values. The mixed effect models were expressed as **Error! Reference source not found.** and 2.

The individual tree basal area growth model was:

$\ln(\gamma_{kij})=X_{kij}b + \mu_k + \varepsilon_{kij}$ *Equation 1*

Where γ_{kij} is logarithm of the individual tree basal area growth of tree *i* in site *k* and period *j* (m^2 $tree^{-1}$ $5\ years^{-1}$), X_{kij} is a vector of the independent variables for tree *i* in site *k* and period *j*; b is a vector of fixed effects, μ_k is random effect of stand *k* and ε_{kij} is residual error for tree *i* in site *k* and period *j*

The per hectare basal area growth model was:

$\ln(\gamma_{kj})=X_{kj}b + \mu_k + \varepsilon_{kj}$ *Equation 2*

Where γ_{kj} is logarithm of the per hectare tree basal area growth of site *k* and period *j* (m^2 ha^{-1} 5 $years^{-1}$), X_{kj} is a vector of the independent variables in site *k* and period *j*; b is a vector of fixed effects, μ_k is random effect of stand *k* and ε_{kj} is residual error for site *k* and period *j*.

With the small dataset that were available for this study, there was unfortunately no possibility to set aside data for validation. This is why the evaluation of the models were made by inspecting residuals following general advice in creating models (Vanclay & Skovsgaard 1997). From the random effect in the model there are many intercepts, however these cannot be accounted for when using the models since they are specific for the random effect of the individual tree or plot. The intercept used will be the mean of all intercepts. In addition to visual inspection of residuals, pseudo-R-square and RMSE values were used as indicators of model performance. The R package MuMIn (Barton 2023) was used to perform a pseudo-R-squared test since a regular R-squared cannot be calculated for mixed effect models. Variables and the transformation of variables were only chosen if they were biologically logical to avoid overfitting and to ensure that the model is based on real biological explanations for growth and development. The models are made with the philosophy that they should strive towards simplicity. The dependent variable was logarithmical transformed, to account for non-linear relationship between the dependent and independent variables. A logarithmical transformation has also been shown to improve basal area models (Liziniewicz et al. 2022). Lastly p-values for the independent variable in the model was also a decisive factor for including variables.

2.1.1 Basal area increment models

For the basal area increment models, BAI1 and BAI2, several independent variables were tested to be included in the model (Table 2). The age of the plots ranged from 17-87 years. The reliability of the models in young stands and very old stands will therefore be unknown due to extrapolation. Some of the independent variables are specific only for BAI1, as they are variables for individual trees.

Site index expressed as top-height at a total age of 100 years (SI), is a measure of fertility and was estimated, to be included in this model. A general height-development functions for larch species in Sweden was used for estimation of SI (Johansson et al. 2013).

Table 2. List of independent variables that was tested in the model for BAI1 and BAI2. BAL and Balgdp1 indicate the tree's position in the stand and act as competition indices, see 2.1.2. The maximum and mean values are only relevant within each stand, where the biggest tree start from the number zero.

Independent variable	Min	Mean	Max	Model
Diameter, cm	0.1	14	50,8	(BA1)
Age, yr	17	32	87	(BA1), (BAI2)
Basal area, m^2 $tree^{-1}$	0.1	0.02	0.2	(BA1)
Basal area, m^2 ha^{-1}	2	19	31	(BA1), (BAI2)
Stems, nr ha^{-1}	167	1048	2391	(BA1), (BAI2)
SI, m	24,1	29.9	34,5	(BA1), (BAI2)
Dominant tree height, m	5,4	14.8	30.8	(BA1), (BAI2)
Latitude	58,3	62.6	64,7	(BA1), (BAI2)
BAL, m^2/ha*				(BA1)
Balgdp1**				(BA1)

* Different variables showing position in a stand. BAL is the tree's basal area relative to other tree's basal area in the stand.

**Balgdp1 is BAL transformed to express growth and competition within the stand.

Of the total number of plots, 46 percent were thinned during the period of measurements. There is a possibility that thinning may influence growth. Generally, thinning will decrease the growth of the stand but increase growth for individual trees in a stand (Gonçalves 2020), but very little for the largest trees (Agestam 2015). Thinning of the experimental plots were done in many ways, with thinning intensities ranging from 10% to 43.5% (Table 3). The difference between mild and moderately intense thinning in basal area growth have shown to be marginal for spruce and pine, after five and ten years (Bianchi et al. 2022). However, a hard thinning showed an effect on basal area growth, which would correspond to the plots with about 40% thinning intensity. There is little evidence for an effect of thinning on top height in Norway spruce and Scots pine for individual trees, if thinned from below (Agestam 2015). The same relationship is assumed for larch in this study. To test the effect of thinning on basal area growth, four different variables were created. For the thinning frequency variable, indicator variables for 0 years, 10 years, and 20 years since thinning as well as thinning intensity, were included in the model testing.

Table 3. Table of independent variables related to thinnings. Limited to removal rate greater than 10%.

Plots: 12

Independent variable	**Min**	**Mean**	**Max**	**Model**
Number of thinnings Number of thinnings within a certain timespan. Tested variables from 1-yrs, 10-yrs and 20-yrs.	1	2,3	4	(BA1), (BAI2)
Thinning strength, % Thinned basal area divided with basal area before thinning.	10%	23%	43,5%	(BA1), (BAI2)
Basal area thinned, m^2 ha^{-1}	1,3	5,2	12,7	(BA1), (BAI2)
Accumulated thinning, m^2 ha^{-1} The sum of thinned basal area.	1,4	11,3	26,9	(BA1), (BAI2)

2.1.2 Variables for estimating competition in the individual tree basal area increment

Basal area of larger trees (BAL) is a measure of the subject trees basal area in relation to other tree's basal area in the stand. This is calculated by summing basal areas of trees larger than the subject tree. The largest tree gets a value of 0, whereas smaller trees get a value of the sum of all basal areas for trees larger than itself. One similar independent variable was also tested to incorporate competition in the model that has been used in other studies, seen in Equation called "Balgdp1" (Elfving 2010). Where "d" stands for diameter of the subject tree at breast height (1.3m above ground).

$$Balgdp1 = \frac{BAL}{(d+1)}$$

Equation 3

2.2 Yield model methodology

The Lundqvist-Korf function (Equation 1) has two parameters that were estimated with the available data. These two parameters were estimated with a "nls" (Nonlinear Least Squares) function in R-studio (Posit team 2023). The model was visually examined for performance. With the two parameters estimated, *b1* and *b2*, and starting values of basal area and total age, the basal area can be estimated at any point in time.

$$y1 = \left(\frac{y0}{b1}\right)^{\left(\frac{t0}{t1}\right)^{b2}}$$

Equation 1

Table 4. Variables for the Lundqvist-Korf formula.

Variable	Explanation
$y1$	Future total basal area
$y0$	Total basal area now
$t1$	Future year
$t0$	Year now
$b1$	Estimated parameter 1
$b2$	Estimated parameter 2

The finished model will be used to calculate the growth for the stands. The residuals for the model can then be analysed and be compared to the BAI1 and BAI2. All plots are made by using the Tidyverse (Wickham et al. 2019) package.

3. Result

3.1 Basal area growth for individual tree

All independent variables used in the model for basal area growth of individual trees were significant (Table 5). The intercept was significant. One thinning variable were significant. The dependent variable was transformed into ln(basal area growth) *m^2 $tree^{-1}$ $5yr^{-1}$*. Five independent variables and two interactions were used in the final model. Residuals for the independent variables can be seen in Appendix 1.

Table 5. Result table for BAI1. Included in the table is RMSE for the tree model and RMSE when the model is converted to a per hectare model. The dependant variable was transformed into ln(basal area growth) m^2 $tree^{-1}$ 5 yr^{-1}. The number of observations is the total amount of repeated measurements and groups are the trees measured.

Number of observations: 6772		Number of groups: 1662	
Variables	Coefficient	P-value	Std.Error
Intercept	-2.2254315	<0.00001	0.09148435
ln(Basal area), m^2 $tree^{-1}$	0.7010126	<0.00001	0.00850789
Age	-0.0095312	<0.00001	0.00136458
BAL, m^2 ha^{-1}	0.0665745	<0.00001	0.00458685
Thin-str, % *	0.0038140	<0.00001	0.00077563
Stems, ha^{-1}	0.0008764	<0.00001	0.00003967
ln(basal area) * BAL	0.0213859	<0.00001	0.00088774
Stems * Age	-0.0000269	<0.00001	0.00000148

* Thinning strength, %

Pseudo-R^2	RMSE	RMSE (ha^{-1})	Degrees of freedom
0.778	0.002079284	1.017489	5053

The BAI1 residuals shows some pattern of heteroscedasticity (Figure 2 and Figure 3). The residual variance grows with increased predicted BA-growth. The residual has been calculated by taking the observed value minus the estimated value. Seeing a positive residual means that the estimated value was lower than the observed. The variance increases more outside the first and third quantile, then inside (Figure 3).

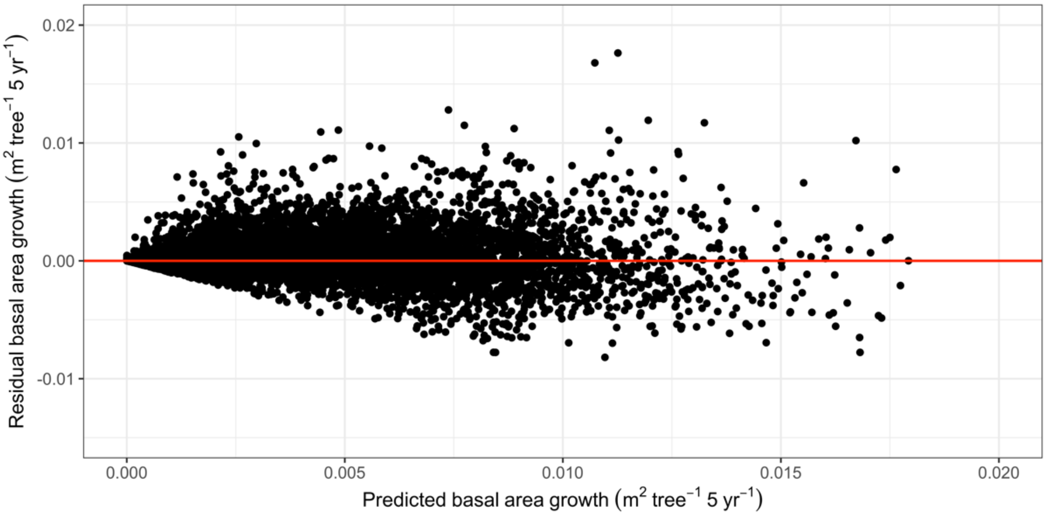

Figure 2. Residual basal area growth against predicted basal area growth for the BAI1 model. The redline is set to zero. There are 6772 observations in this model.

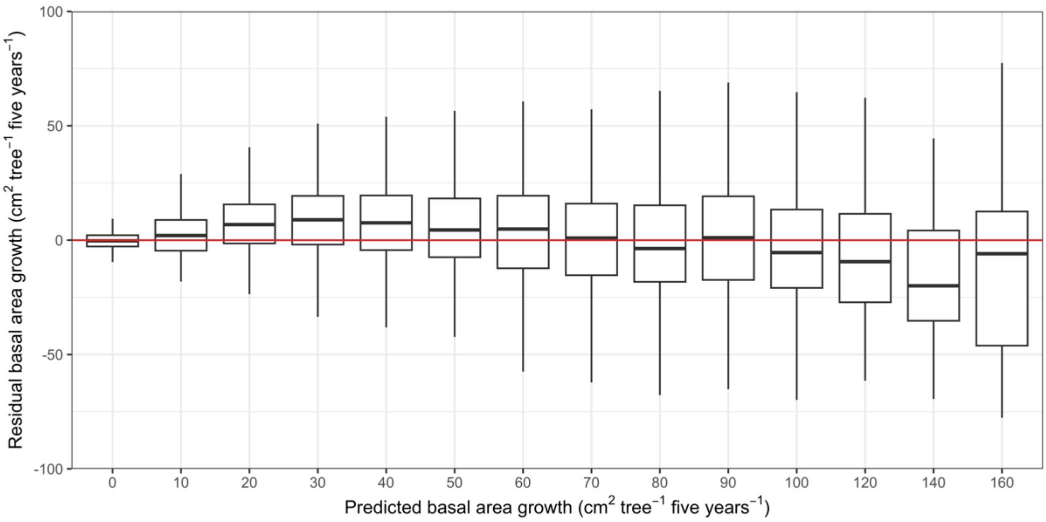

Figure 3. Boxplot over the residual variance in cm^2 $tree^{-1}$ 5 yr^{-1} for the BAI1 model. Note from figure 2 that most of the observations are between 0-100 cm^2.

The residuals for the converted BAI1 to the per hectare model, can be seen in Appendix 4. It shows some bias and a more notably heteroscedastic relationship to the predicted variables than before the conversion.

3.2 Basal area growth per hectare

The variables that were chosen for the BAI2-model were on stand level which means the number of observations were much lower than for the individual tree model. All variables used in the model were significant, but the intercept was not (Table 6). The high Pseudo-R^2 and the relatively low RMSE indicate that the model performs well. Residuals for the independent variables can be seen in Appendix 1.

Table 6. Result table for BAI2. Note that the Pseudo-R^2 is higher than 1. The dependant variable was transformed into ln(basal area growth) m^2 5 yr^{-1} ha^{-1}. The number of observations is the total amount of repeated measurements and groups are the plots measured.

Number of observations: 76		Number of groups: 26	
Variable	Coefficient	P-value	Std.Error
Intercept	0.1751528	0.5311	0.27746991
ln(Basal area), m^2 ha^{-1}	0.6725128	<0.00001	0.07994983
Age	-0.0204937	<0.00001	0.00335542
Stems, ha^{-1}	0.0008326	<0.00001	0.00011644
Stems * Age	-0.0000261	<0.00001	0.00000503
Pseudo-R^2	RMSE	Degrees of freedom	
1.001605	0.8561552	44	

Residual versus predicted basal area growth follow a 1:1 line relatively well (Figure 4). Lower accuracy in the model was in the higher growths. The variance is at its highest at 7.5 *m^2 5 yr^{-1} ha^{-1}*, it is however lower in its residual variance after that point. No bias can be seen in the residuals, and there is no clear pattern either. A residual variance boxplot can be seen in Appendix 3.

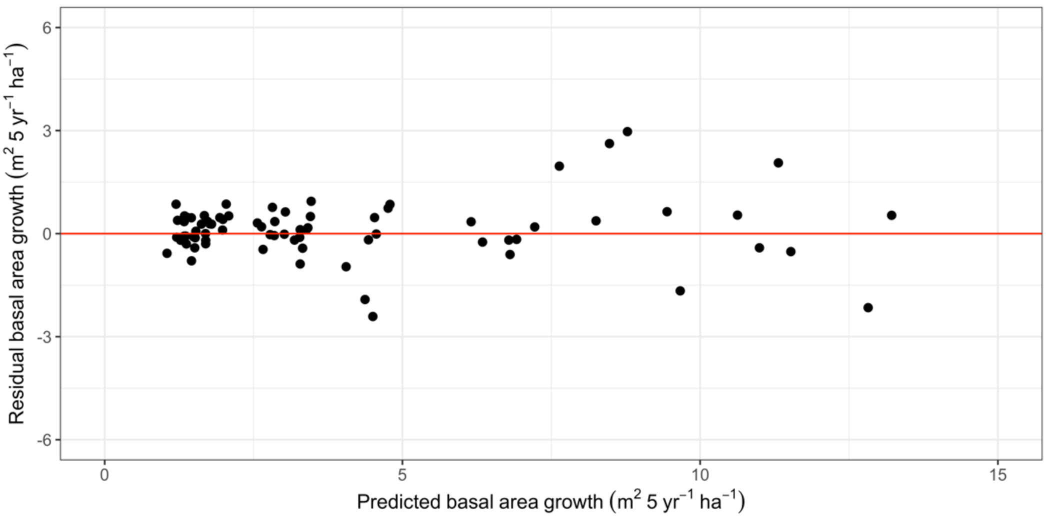

Figure 4. Residuals for BAI2 model. Note that the y-axle has a different relation of scale compared to figure 2.

3.3 Basal area yield per hectare

The two estimated parameters from the *nls* function, $b1$ = 73.41482 and $b2$ = 1.33154. A visual inspection of the model shows a slight underestimation of high values and an overestimation of low values (Figure 5) High SI will give greater total basal area yield, even if it was not part of the model per se.

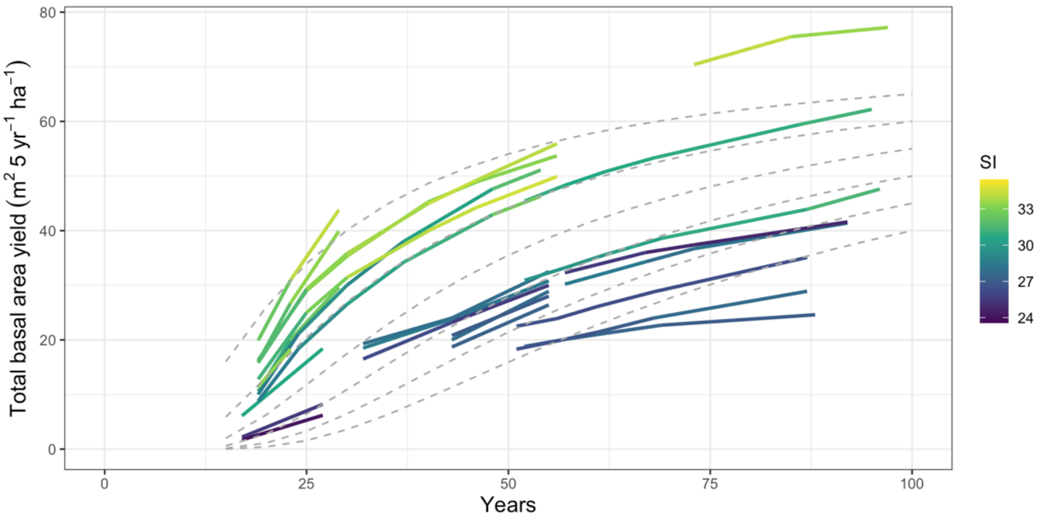

Figure 5. Total basal area yield with measured data and the yield model fitted. The measured data with coloured based on SI and the model fitted as grey dashed lines. The models have six lines that are calculated total basal area yield at 100 years, starting from 40 m^2 and increasing with 5 m^2 until 65 m^2. The lines are calculated backwards in 5-year increments to year 15.

The yield model was converted to a growth model (BAI3), and the residuals created put against the measured data for BA-growth m^2 ha^{-1}, as seen in Figure 6. The model underestimates in the higher growths and overestimates in the lower. This model has a RMSE = 1.735734. The predicted growth does not go higher than 10 m^2 5 yr^{-1} ha^{-1}, while there are around 10 plots that is higher than that. The residual variance can be seen in Appendix 3.

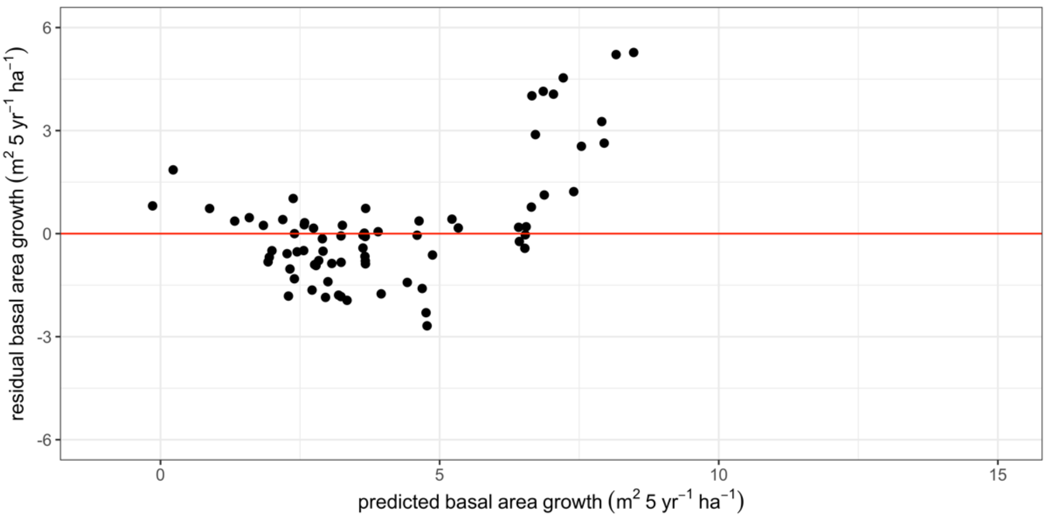

Figure 6. Residuals for model BAI3, which was derived from the BAY model, against the measured basal area growth.

3.4 Comparing all Models for growth and yield

The three models show some similar pattern to the measured data in Figure 7. BAI2 shows follow the 1:1 ratio the best and has the lowest RMSE. BAI3 show a clear underestimation for high growth-values. BAI2 underestimates slightly in the higher growths.

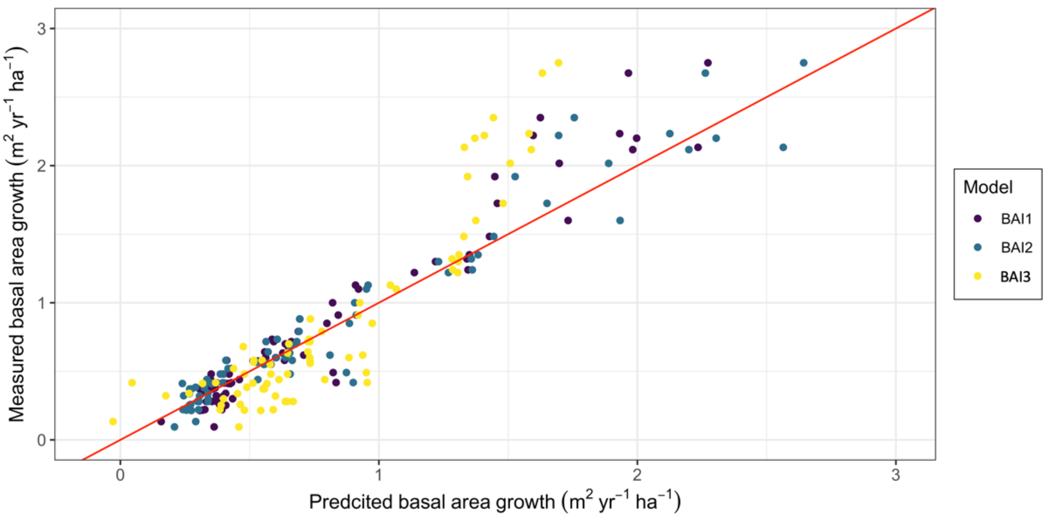

Figure 7. Measured basal area growth from the three models against predicted basal area growth. The models BAI1 and BAI3 were converted into a growth model per hectare for this comparison. There are 74 values for each model.

4. Discussion

4.1 Evaluation

4.1.1 Growth models

The pseudo-R^2 seem to increase when transforming variables for the models. Therefore, it was used mostly as a reference of change for which variables could be included, but not when looking at different transformation. When determining which transformations should be used, a visual inspection of the residuals and RMSE was important. The pseudo-R^2 value for BAI2, with an unexpected value above 1 showed that pseudo-R^2 may be misleading. The intercepts for BAI2 were not statistically different from zero as the p-value was higher than 0.05. Therefore, the intercept should do not need to be included in those models.

BAI1 shows some level of heteroskedasticity with higher growths. The amount of data in the lower growth poses a problem with an uneven distribution of the dependent variable. The median growth for individual trees of the measured data is 46 $cm^2\ tree^{-1}\ 5\ yr^{-1}$. The dependant variable was chosen to be logarithmical transformed, which lowers the heteroskedasticity, but there is some possibility that a higher logistical base could be chosen for the benefit of the model. When BAI1 is converted to a per hectare model, a bias and clearer heteroskedasticity can be observed.

The models have a group of overestimations around a measured basal area growth of 3 $m^2\ 5\ yr^{-1}\ ha^{-1}$. They are from the same site and first revision, (see appendix 2). Interestingly they were overestimated by all the models. There could be a strong random effect there for which the BA-growth models cannot predict. BAI1 and BAI3 also overestimates the second revision, which BAI2 did not. Another "problematic" area with the name of "Renberget" had an extraordinary growth, which gives the highest absolute value residuals for the models. BAI2 had the lowest residual value and seem to have little systematic error in terms of high growth. This shows that the potential random effect from plots and individual trees could have on the models. This is hard to counteract when predicting with the help of the models and will be part of the error.

Many combinations of independent variables were tested and analyzed in the creation of these models. For the BAI1 model 5 variables and two interactions were included. This could create an over-fitting scenario, making the model predictions less accurate when used for other data.

A slight collinearity problem can be observed between age and number of stands. This collinearity problem makes sense, when a stand gets older and is managed with a classic silviculture system, the number of stems will be much lower in the end of the rotation than in the beginning. The same goes for basal area for the individual trees, most of the smaller ones have been thinned. These relationships could cause a problem when using a model on new stands that have another ratio of age, basal area and stems per hectare. Especially when using the BAI2 which is only utilizing these tree independent variables.

The final models chosen only included significant independent variables, but reasons for being significant could be due to random events or being biased. This is hard to investigate without validating the model on an independent dataset of other Siberian larch stands in Sweden or neighboring countries.

4.1.2 The growth models regarding thinning variables.

The lack of thinning variables in BAI2 means that the potential effect of thinning was not accounted for. An unthinned stand will have higher frequency of smaller diameter than a stand that has gone through a thinning regime, since the thinnings were assumed to be done from below (Agestam 2015). The plots used in these models have been thinned 0-5 times. One probable reason for the lack of a significant thinning variable, could be that there were too few unthinned old stand. There was no thinning-experiment with an unthinned control included in the data. The unthinned observations were mostly found in young where the lack of thinning may not yet have had a significant effect on growth on the whole stand. However, there is a possibility that the other independent variables explain enough of the effect that the thinning would have on the stand, for example basal area and stems per hectare after thinning and age, and their interactive effect.

Depending on how many thinnings a stand had undergone, the retained trees could have been affected with increased growth, compared to trees in potential unthinned stands. Guidelines regarding the thinning regime of Siberian larch shows that its needed to be thinned usually up to 4-5 times with short intervals (Arvidsson 2006). If not managed correctly the crowns could become too small creating a lower production when the stand later is thinned. This loss of production that could be caused by to late thinnings, is probably not possible for any of the models to consider accurately.

4.2 Yield model

BAY model utilizes only age and total basal area. This means that all previous thinning's basal area should be known, which makes it the only variable from all models that needs a recorded history of measurements from previous forest measures. The yield model will probably not give a great prediction if used on young stands. Converting it to a growth model shows that it is far from being useful to estimate the future growth. There are many other models that can be tested for estimating total basal area production and it is possible that a model with better fit can be found. However, this was beyond the scope of this study.

4.3 Use of models

Of the three models for basal area growth, BAI2 performed best. BAI2 had the lowest RMSE, highest pseudo-R^2, few systematic errors and had the least problem with high growth stands. BAI2 utilizes only a few variables that are easy to measure, making it preferable from a usage point of view. The variable chosen in the final model seems to have limited the underestimations of high growth, but there is still chance that the model when used on new fast-growing stands will underestimate growth. Comparing to the model used today in the Heureka DSS shown in Carlzsons work (2022), the estimated growth $m^2\ yr^{-1}\ ha^{-1}$ has the same pattern of error as BAI3, (Figure 8). The prediction in that work have used almost the same data as in this study and the problematic plots for the Heureka models were the same ones as for the models from this study. The residuals of the Heureka model have a higher residual value then the models created in this study, especially compared to the BAI2 model.

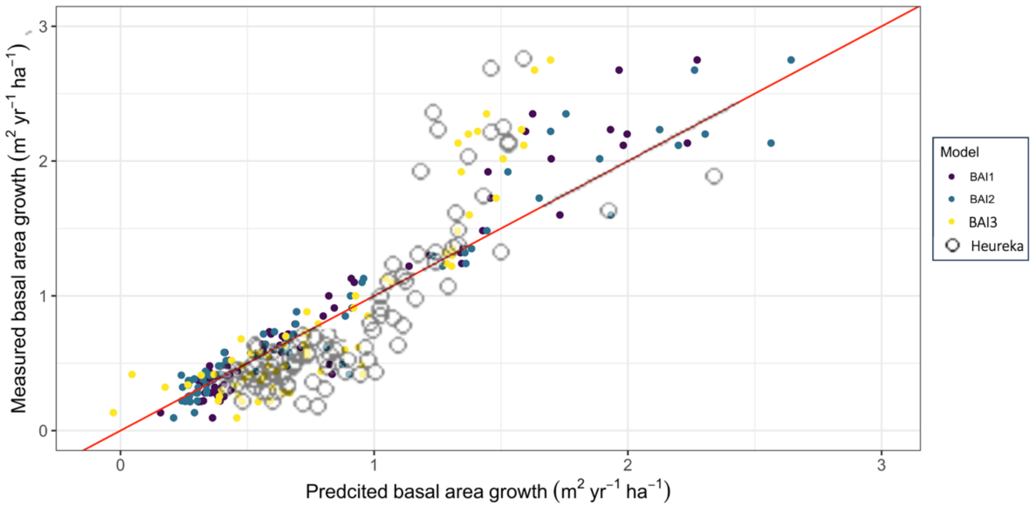

Figure 8. Overlapping figure 7 with figure 4 (lower figure) from Carlzons (2022) work. Showing the difference in predicting the growth compared to that of Heureka. Hollow circular points are from the predictions made with Heureka DSS.

Growth in the Heureka model correlated to stand age shows systematic error pattern of under- and overestimations from age 20-70 years, (Figure 9). From a visual observation it looks like the Heureka model cannot consider the early high growth Siberian larch stands are producing. It also seems to be overestimating between 30 years and 70 years. This makes sense as siberian larch has a high growth rate in early ages that seem to stagnate quick, in relation to age. Making it more similarly to a fast-growing exotic pioneer species, then to Scots pine.

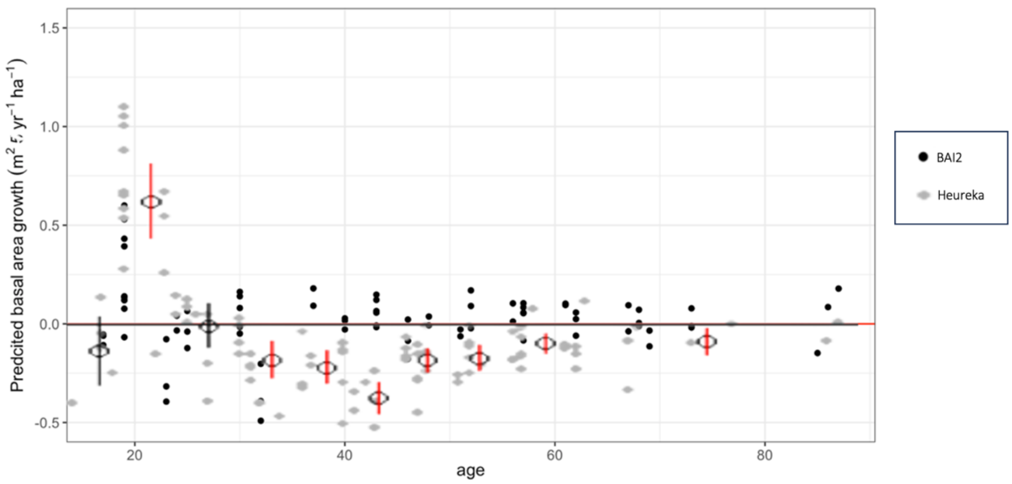

Figure 9. Predicted BA-growth against age with an overlap of plot figure 6 (upper figure) from Carlzons 2022 work. Showing the difference in predicting the growth compared to that of Heureka. Grey dots and hollow circular points with "whiskers" are from Carlzons work.

4.4 Further Research

There is still a lot of unknowns to Siberian larches stand development. What could be of most interest is to create an accurate volume model, that could together with BAI2, be incorporated into the decision support system Heureka. This would enable users to simulate and plan the potential of Siberian larch with more accuracy and compare it to other alternative species. The BAI2 model also need to be validated against an independent data-set.

Conclusion

Out of the three models created BAI2 estimated growth the best. It shows the least amount of RMSE and systematic error. It utilizes three independent variable and one interaction. Compared to BAI1 and BAI3, the BAI2 model does not show a clear systematic pattern of underestimation for higher growths. BAI2 have some problem with underestimation and heteroskedasticity in the stands with high growth.

The BAI2 model is at this moment a good option to predict the growth of Siberian larch stands. Compared to models currently used in the Heureka decision system, BAI2 perform much better with smaller residuals. However, data for estimating the coefficients for BAI2 were relatively limited and should be tested against an independent data before being used in Heureka. The limited amount of data also means that the models should be used with caution, especially when used at the lower- and higher ends of the data-set.

References

Abaimov, A.P., Lesinski, J.A., Martinsson, O. & Milyutin, L.I. (1998). *Variability and ecology of Siberian larch species*. (Rapporter, 43). Umeå: Swedish University of Agricultural Sciences -Department of Silviculture.

Agestam, E. (2015). *Skogsskötselserien - Gallring*. andra upplagan. www.skogsstyrelsen.se/skogsskotselserien [2023-01-13]

Arvidsson, B. (2006). *Sibirisk Lärk*. (Faktablad från Svenska skogsplantor). Svenska skogsplantor. https://www.skogskunskap.se/cd_20180103095409/contentassets/087017b8a7c34dd0a6a1661a643ebb5d/sibirisk-lark-2011-svenska-skogsplantor.pdf

Assmann, E. (1970). Introduction. In: *The Principles of Forest Yield Study*. München: Pergamon. 1–5. https://doi.org/10.1016/B978-0-08-006658-5.50004-2. [2023-03-15]

Barton, K. (2023). *Multi-Model Inference* (1.47.5) [R (≥ 4.2.0)]. https://cran.r-project.org/web/packages/MuMIn/index.html

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67 (1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bianchi, S., Huuskonen, S., Hynynen, J., Siipilehto, J. & Niemistö, P. (2022). Tree-level differences in Norway spruce and Scots pine growth after extreme thinning treatments. *Scandinavian Journal of Forest Research*, 37 (2), 109–118. https://doi.org/10.1080/02827581.2022.2045348

Carlzon, S. (2022). *Validation of basal area growth functions for larch in Heureka DSS*. (Avancerad nivå, A2E). SLU, Institutionen för sydsvensk skogsvetenskap. https://stud.epsilon.slu.se/18219/ [2022-10-08]

Elfving, B. (2010). *Growth modelling in the Heureka system*. Umeå: SLU. https://www.heurekaslu.se/w/images/9/93/Heureka_prognossystem_%28Elfving_rapportutkast%29.pdf [2023-01-30]

Feng-ri, L., Bao-dong, Z. & Gui-lin, S. (2000). A derivation of the generalized Korf growth equation and its application. *Journal of Forestry Research*, 11 (2), 81–88. https://doi.org/10.1007/BF02856679

Gonçalves, A.C. (2020). *Thinning: An Overview*. *Silviculture* IntechOpen. https://doi.org/10.5772/intechopen.93436

Goude, M., Nilsson, U., Mason, E. & Vico, G. (2022). Comparing basal area growth models for Norway spruce and Scots pine dominated stands. *Silva fennica (Helsinki, Finland : 1967)*, 56 (2). https://doi.org/10.14214/sf.10707

Heiðarsson, L. & Pukkala, T. (2012). Models for simulating the temporal development of Siberian larch (Larix sibirica) plantations in Hallormsstaður Iceland. *ias*, 25, 13–23

Johansson, U., Ekö, P.M., Elfving, B., Johansson, T. & Nilsson, U. (2013). Nya höjdutvecklingskurvor för bonitering. *FAKTA SKOG*, 14

Karlman, L. (2010). *Genetic variation in frost tolerance, juvenile growth and timber production in Russian larches (Larix Mill.): implications for use in Sweden*. Department of Forest Ecology and Management, Swedish University of Agricultural Sciences.

Korf, V. (1939). *A mathematical definition of stand volume growth law*. (Lesnicka Prace 18)
Kullman, L. (1998). Palaeoecological, Biogeographical and Palaeoclimatological Implications of Early Holocene Immigration of Larix sibirica Ledeb. into the Scandes Mountains, Sweden. *Global Ecology and Biogeography Letters*, 7 (3), 181–188. https://doi.org/10.2307/2997373
Liziniewicz, M., Barbeito, I., Zvirgzdins, A., Stener, L.-G., Niemistö, P., Fahlvik, N., Johansson, U., Karlsson, B. & Nilsson, U. (2022). Production of genetically improved silver birch plantations in southern and central Sweden. *Silva Fennica*, 56 (1). https://doi.org/10.14214/sf.10512
Lundqvist, B. (1957). *Om höjdutvecklingen i kulturbestånd av tall och gran i Norrland On the height growth in cultivated stands of pine and spruce in northern Sweden*. Stockholm: Statens skogsforskningsinstitut. (Meddelanden från Statens skogsforskningsinstitut, 47:2)
Martinsson, O. (1995). Yield of Larix suhaczewii Dyl. in NorthernSweden. *Studia Forestalia Suecica*, No. 196
Nilsson, P., Roberge, C. & Fridman, J. (2021). *Skogsdata 2021*. Petersson, H. (ed.) (Petersson, H., ed.) 400 ex. Umeå: SLU Institutionen för skoglig resurshushållning.
Norén, M. (2009). *Regler om användning av främmande trädslag*. (7). Jönköping: Skogsstyrelsen.
Normark, E. & Fries, C. (2019). *Skogsskötsel med nya möjligheter*. (Samverkansprocess skogsproduktion, 2019/24). Jönköping: Skogsstyrelsen. https://www.skogsindustrierna.se/siteassets/dokument/nyheter/rapport-fran-samverkansprocess-skogsproduktion.pdf
Panik, M.J. (2013). Parametric Growth Curve Modeling. In: *Growth Curve Modeling*. United States: John Wiley & Sons, Incorporated. 49–87. https://ebookcentral.proquest.com/lib/slub-ebooks/detail.action?docID=1577052. [2023-03-06]
Pesonen, A., Eerikäinen, K., Maltamo, M. & Tahvanainen, T. (2009). Models for predicting tree and stand development on larch plantations in Hallormsstaður, Iceland. *New Forests*, 37 (1), 63–83. https://doi.org/10.1007/s11056-008-9107-6
Posit team (2023). *RStudio: Integrated Development Environment for R* (2022.07.1+554.pro3 + 2022.12.0.353). Boston, MA: Posit Software, PBC. http://www.posit.co/.
Sedmák, R. & Scheer, L. (2015). Properties and prediction accuracy of a sigmoid function of time-determinate growth. *iForest - Biogeosciences and Forestry*, 8 (5), 631–637. https://doi.org/10.3832/ifor1243-007
Söderholm, E. & Öhman, M. (2010). *Potentiella arealer för ökad virkesproduktion i norra Sverige genom skogsodling med Pinus contorta var. latifolia, Larix sukaczewii, Abies lasiocarpa, Populus × wettsteinii och Picea mariana*. (Grundnivå, G2E). SLU, Institutionen för skogens ekologi och skötsel. https://stud.epsilon.slu.se/2016/ [2022-10-01]
Valinger, E. & Fridman, J. (2011). Factors affecting the probability of windthrow at stand level as a result of Gudrun winter storm in southern Sweden. *Forest Ecology and Management*, 262 (3), 398–403. https://doi.org/10.1016/j.foreco.2011.04.004
Vanclay, J.K. & Skovsgaard, J.P. (1997). Evaluating forest growth models. *Ecological Modelling*, 98 (1), 1–12. https://doi.org/10.1016/S0304-3800(96)01932-1
Weiskittel, A.R., Hann, D.W., Kershaw, J.A.Jr. & Vanclay, J.K. (2011). *Forest growth and yield modeling*. 1. ed. New York: John Wiley & Sons, Incorporated.

Westin, J., Helmersson, A. & Stener, L.-G. (2016). *Förädling av lärk i Sverige – Kunskapsläge och material*. (Arbetsrapport, 843). Uppsala: Skogforsk. https://www.skogforsk.se/cd_20190114161536/contentassets/e077be5a553043759819a5937d1be289/lark--kunskapslage-material-och-foradlingsmojligheter-arbetsrapport-895-2016.pdf

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4 (43), 1686. https://doi.org/10.21105/joss.01686

Predicting the future in Siberian larch forests

- *Creating new Growth and Yield models for Siberian larch forests.*

One common question forest manager asks themselves are: how much will the forest grow in the next decade? This information is needed when evaluating which forest management actions should be taken and when interventions should be done. With the help of growth and yield models, these estimations can be made.

One previous study found that the model used for predicting the growth of Siberian larch in Sweden was inaccurately estimating basal area[1] growth compared to real measured trees. Siberian larch counts as a native tree species in Sweden, which makes it one of three native coniferous tree species suitable for production forestry. With less than 0.1% of Sweden consisting of Siberian larch there is little accurate information of its growth properties.

Three different methods were tried to create new models that could accurately predict its growth. One new model was then selected among the candidates which showed clear improvement. The new model can predict more accurately the future growth of Siberian larch forests.

If you want to know how much your Siberian larch forest will grow in the next 5 years, all you need is age, number of stems ha^{-1} and the current basal area $m^2\ ha^{-1}$.

$$\textit{Basal area growth} = \exp(\ln(\text{Basal area}) * 0.6725128 + \text{Age} * -0.0204937 + \text{Stems} * 0.0008326 + \text{Stems} * \text{Age} * -0.0000261)$$

[1] Basal area per hectare is the total cross-sectional area of trees at 1.3m above the ground level.

Acknowledgements

I would like to express my gratitude to my supervisor Urban, who has been helping me throughout this paper. He has not only guided me throughout the process but has also been a constant source of help and insightful advice.

Appendix 1

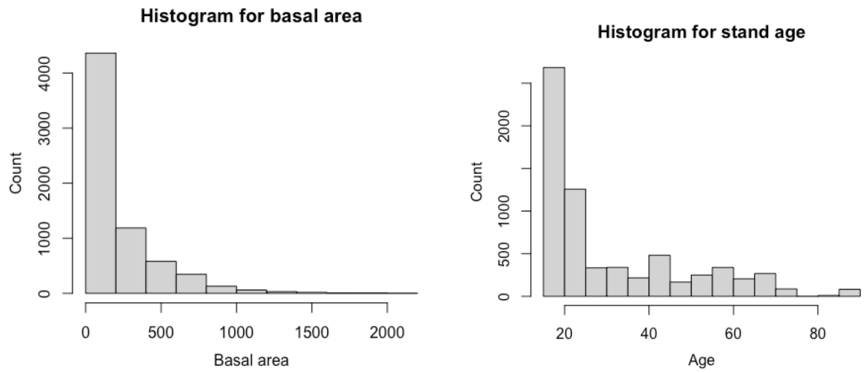

Figure 1. Histogram of basal area, left figure, and histogram of age, right figure, for model BAI1.

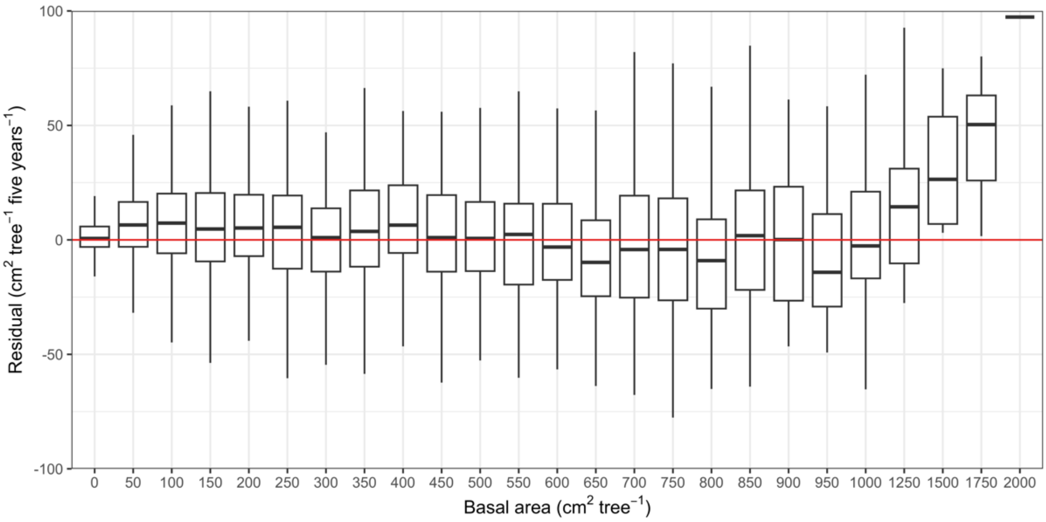

Figure 2. Residual variance for basal area for model BAI1.

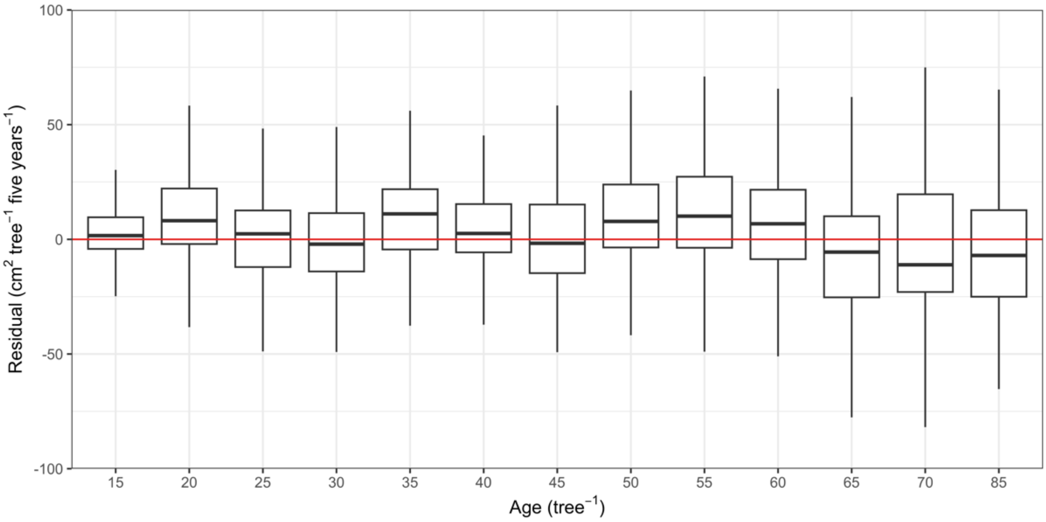

Figure 3. Residual variance for Stand age for model BAI1.

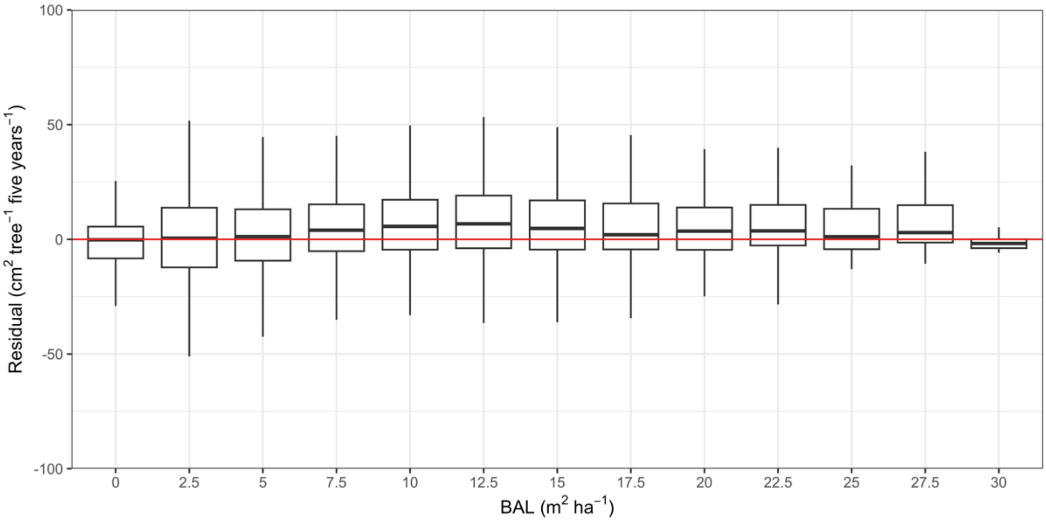

Figure 4. Residual variance for BAL for model BAI1.

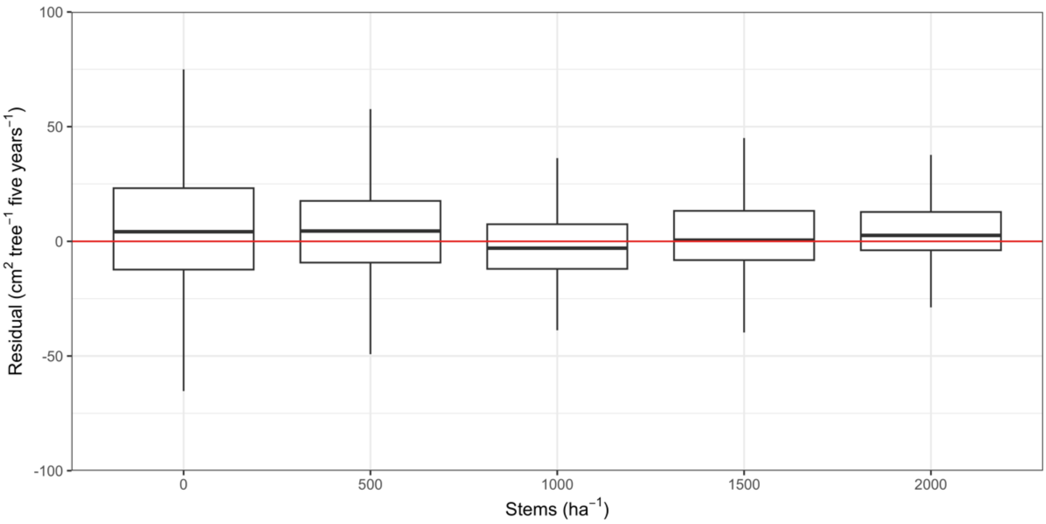

Figure 5. Residual variance for Stems per hectare for model BAI1.

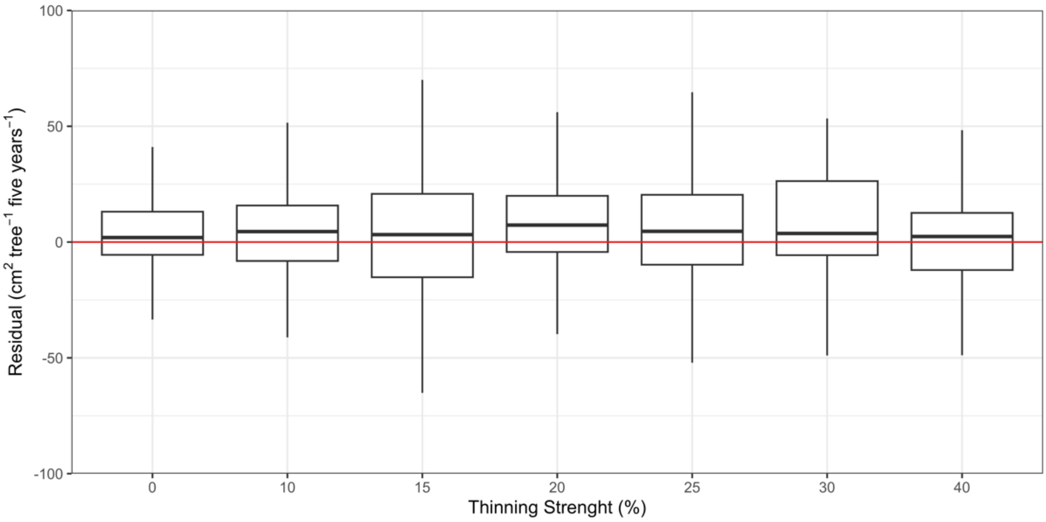

Figure 6. Residual variance for Thinning strength for model BAI1.

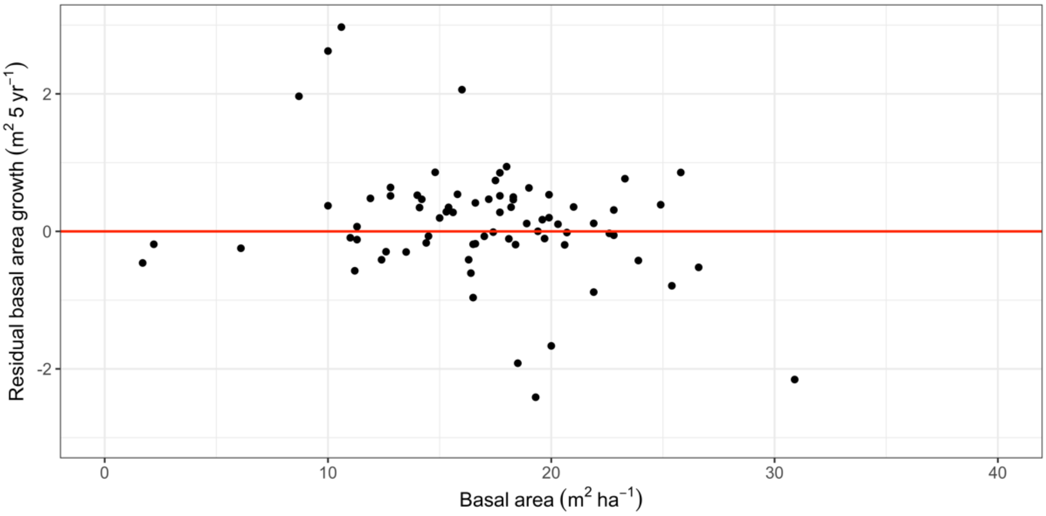

Figure 7. Residuals for Basal area per hectare for model BAI2.

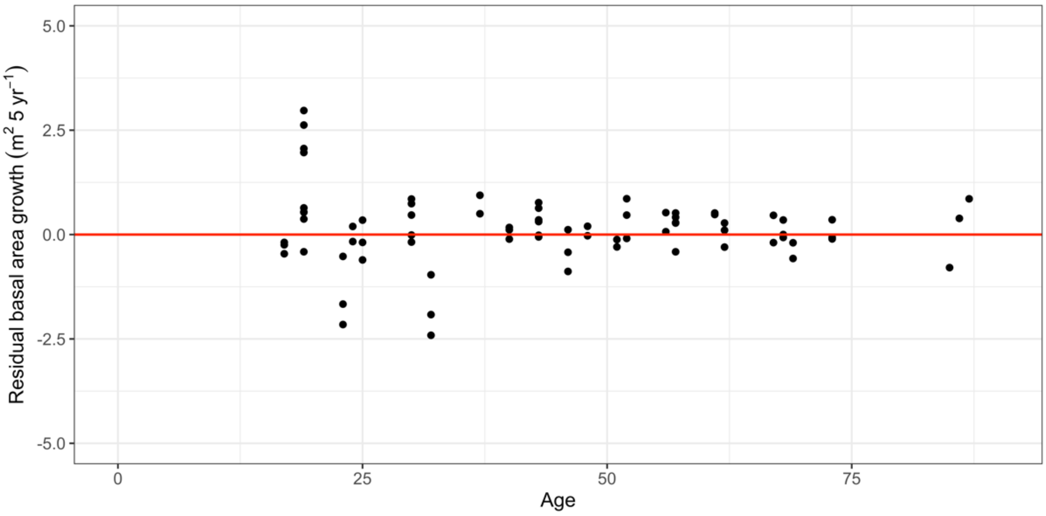

Figure 8. Residuals stand age for model BAI2.

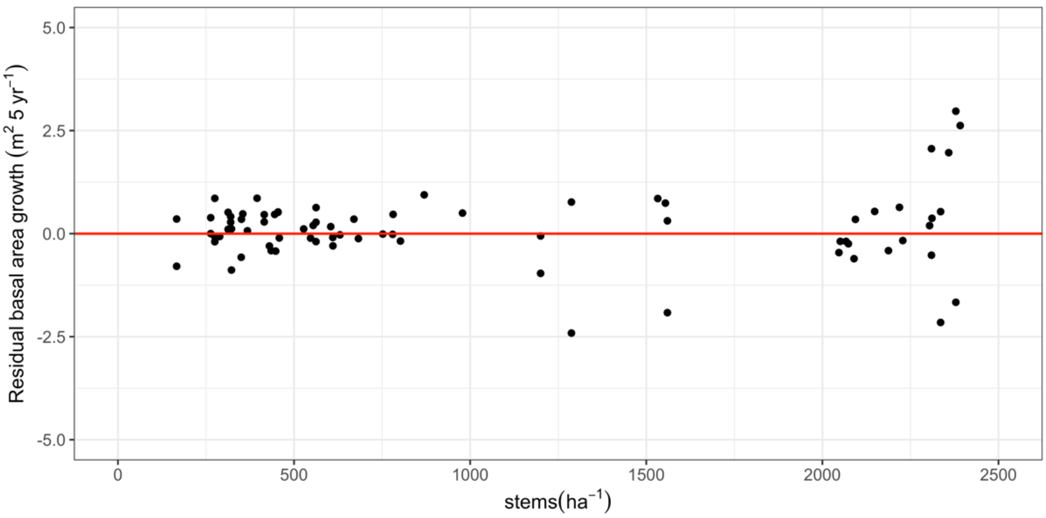

Figure 8. Residuals stems per hectare for model BAI2.

Appendix 2

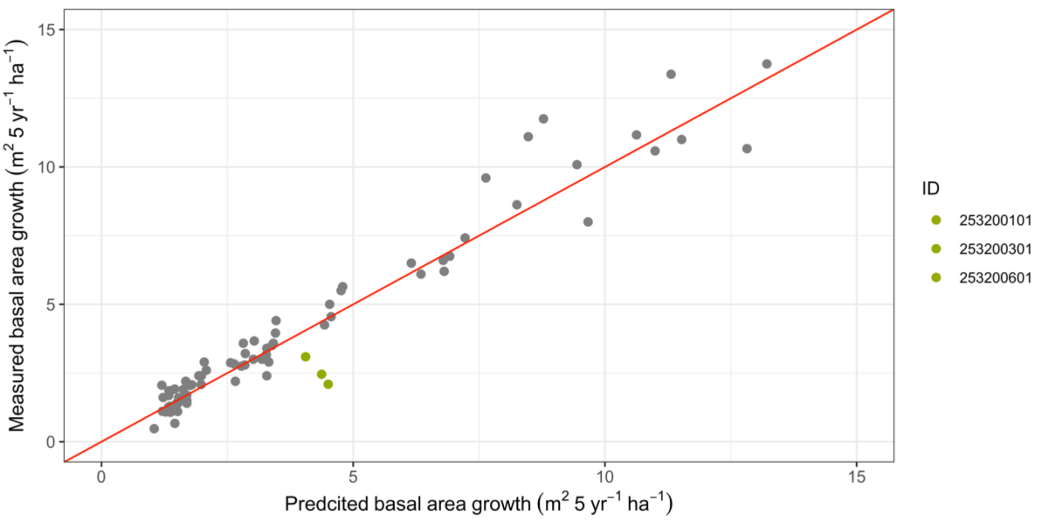

Figure 1. BAI2 with highlight on three plots in a site. They are all the revision one stands and all three of them and are standing out. The model is overestimating them greatly.

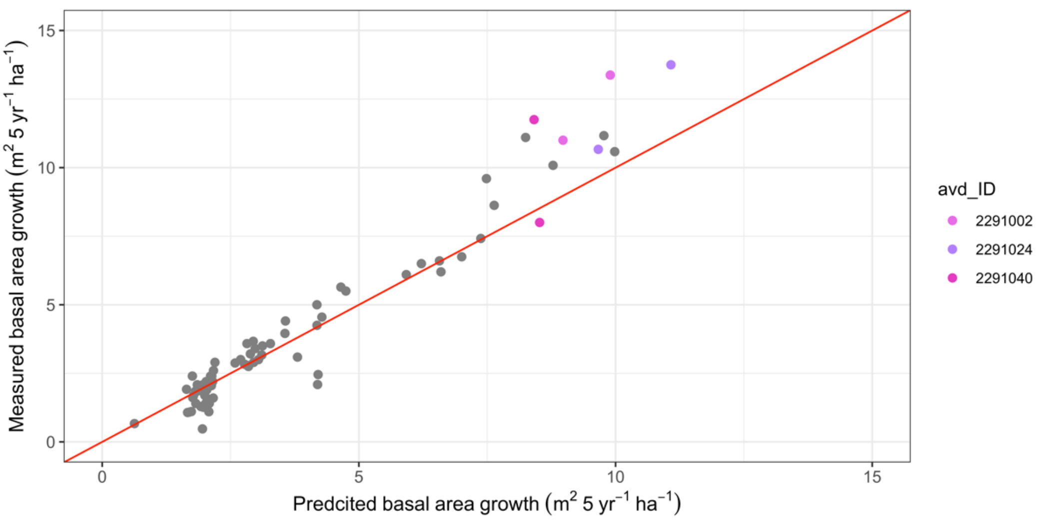

Figure 2. BAI1 converted into growth per hectare. 2291 is from a place called Renberget which has shown outstanding growth and the model could not take this into account. A systematic error can be observed.

Appendix 3

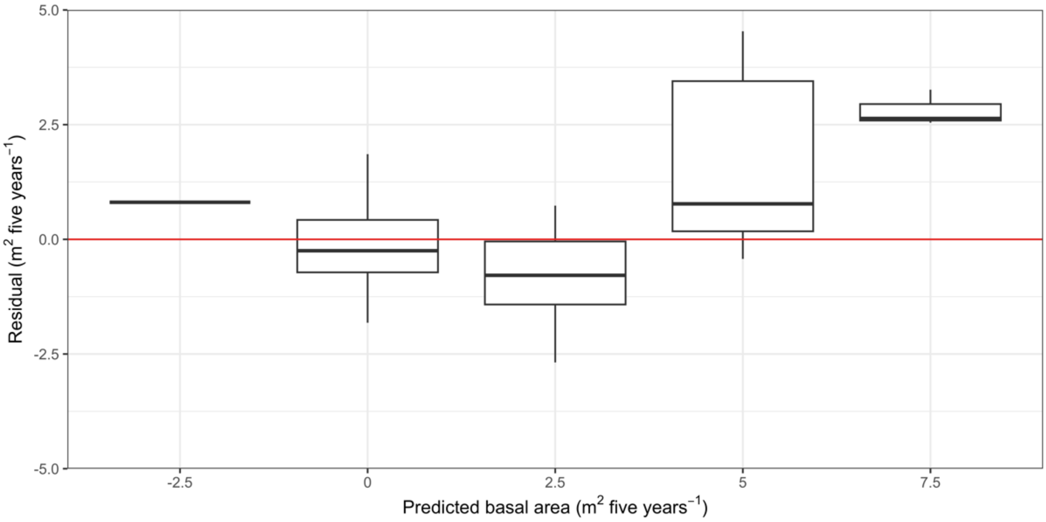

Figure 1. Residual variance of the BAI3 model.

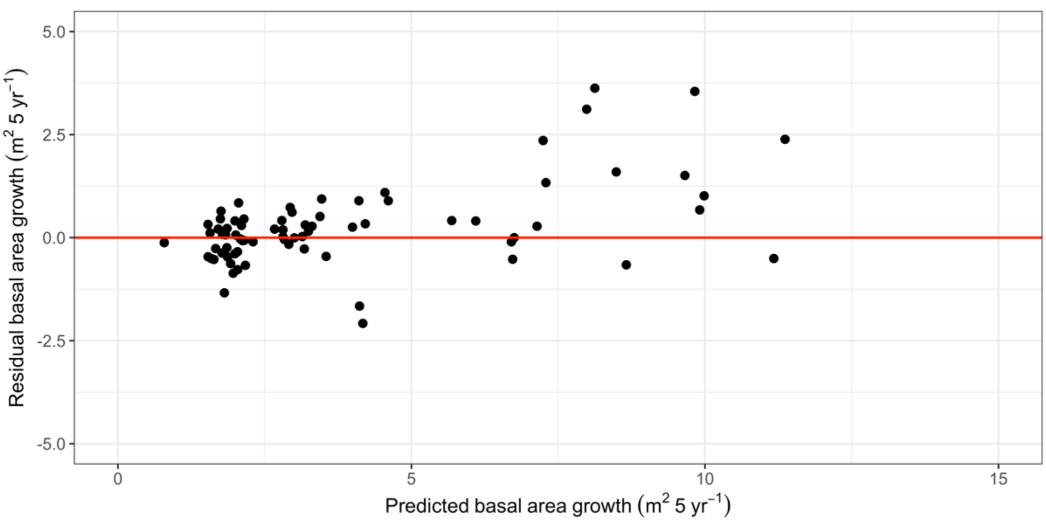

Figure 2. Residual plot for BAI1 converted to per hectare model.

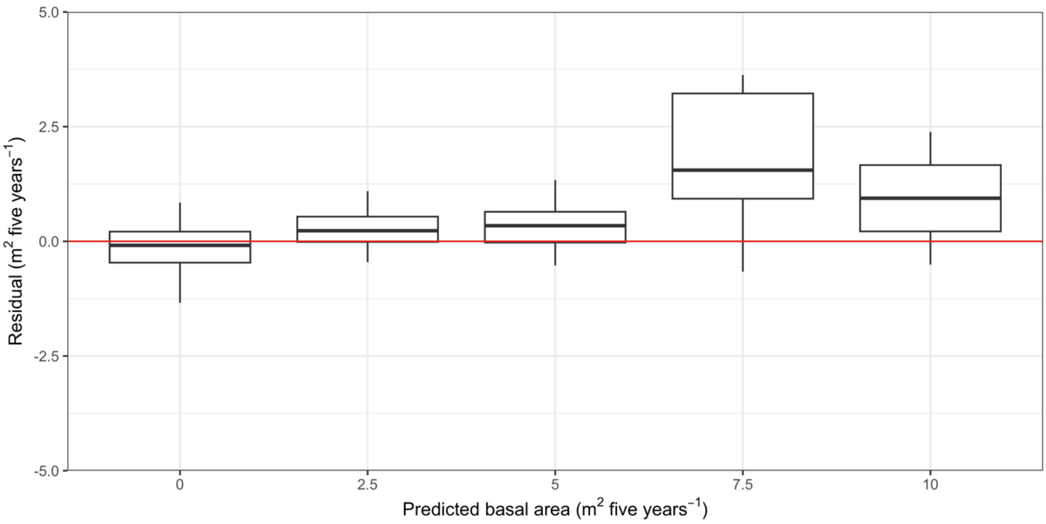

Figure 3. Residual variance for BAI1 converted to per hectare model.

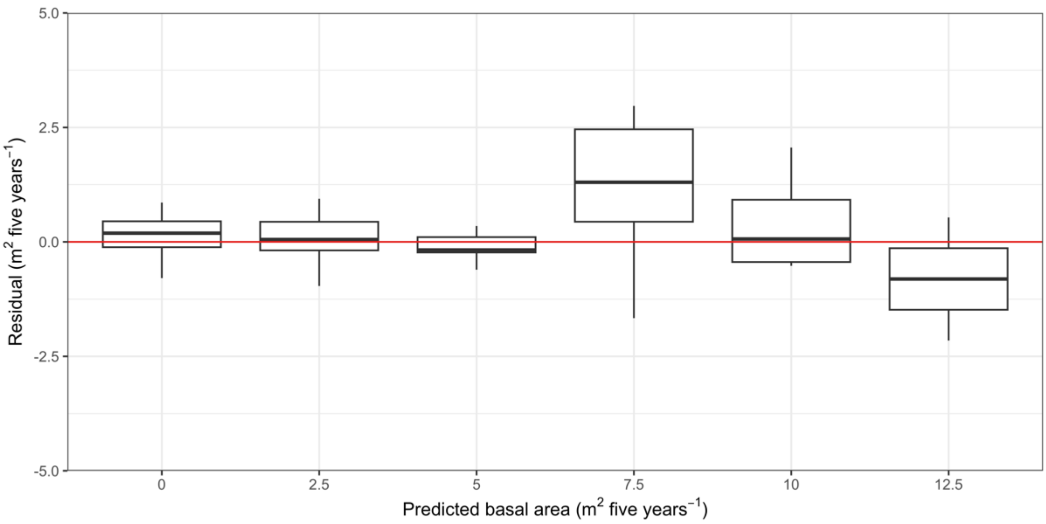

Figure 4.. Boxplot of residual variance for the BAI2 model. Note that the Y-axle and X-axle is in a different scale then.

Publishing and archiving

Approved students' theses at SLU are published electronically. As a student, you have the copyright to your own work and need to approve the electronic publishing. If you check the box for **YES**, the full text (pdf file) and metadata will be visible and searchable online. If you check the box for **NO**, only the metadata and the abstract will be visible and searchable online. Nevertheless, when the document is uploaded it will still be archived as a digital file. If you are more than one author, the checked box will be applied to all authors. You will find a link to SLU's publishing agreement here:

- https://libanswers.slu.se/en/faq/228318.

☒ YES, I/we hereby give permission to publish the present thesis in accordance with the SLU agreement regarding the transfer of the right to publish a work.

☐ NO, I/we do not give permission to publish the present work. The work will still be archived and its metadata and abstract will be visible and searchable.