Characterizing dependence of Irish Sitka spruce stands using spatio-temporal sum-metric models

Abstract

Individual tree dependence in forest plots is spatially dependent, changes over time and the magnitude of spatial dependence may also change over time, particularly in stands subjected to thinning. Models for tree dependence in the literature have been mainly restricted to either spatial models or temporal models. We extend these to spatio-temporal models. The data are from three long-term, repeatedly measured, experimental plots of Sitka spruce (*Picea sitchensis* (Bong.) Carr.) in Co. Wicklow, Ireland with thinning treatments of unthinned, 40% thinned, and 50% thinned, respectively. A model for tree diameter at breast height, over all locations in each plot and all time points, was fitted with fixed covariates and with a sum-metric spatio-temporal variogram for the covariance structure. In the variogram, the spatial correlation component followed a wave function (due to competition at small distances). The correlation over time also followed a wave variogram while the spatio-temporal anisotropy captured the space-time interaction. The models indicate, once fixed effects are accounted for, that spatial variability and correlation is more important than temporal. Models were fitted to plots with three different treatments to demonstrate model parameters differed by thinning type but were consistent in their interpretation with thinning type. The models show that describing spatial dependence is important in understanding the nature of tree growth and its prediction.

KEYWORDS: Negative autocorrelation; regression-kriging; spatio-temporal tree interactions; wave covariance function.

1. INTRODUCTION

Sitka spruce is the main tree species in Irish forests accounting for approximately 52% of total forest area (Forest Service 2007). Understanding and modeling its growth pattern is a key component in making informed management decisions such as when to thin and when to clearfell. The main models used for estimating individual tree volume for commercial tree species, including Sitka spruce, in Ireland, are the British Forestry Commission single tree tariff chart models (Matthews et al. 2006), and the dynamic growth and yield stand models for the main commercial coniferous tree species in Ireland developed by Broad and Lynch (2006). Neither of these models have an explicit spatial component.

Fox et al. (2007*a*) discusses how the growth of individual trees are subject to interacting influences, two of which are competition and micro-site variability. Competition can be defined as the negative effect of one tree on another by consuming, or controlling access to, a resource that is limited in availability (Keddy 1989), while micro-site variability describes local, spatial variation in soil, topographic, geologic and micrometeorological factors (Matern 1960). Competition tends to create negative dependence in size or growth of trees in spatial proximity. In a subsequent paper, Fox et al. (2007*b*) used ARMA models and direct specification of the spatial covariance function to model spatial dependence. However, the residuals from the model exhibited negative spatial dependence over short inter-tree distances. Schabenberger and Gotway (2005) describe the shortcomings of analyzing spatial and temporal effects as separate two-stage approaches.

For a natural forest the locations of the trees are a point process and some characteristic of the plants (e.g. diameter) is the mark (Pommerening and Särkkä 2013). In a spatial analysis, Pommerening and Särkkä (2013) showed mark variograms of incremental tree diameters of Norway spruce (Austria) exhibited negative autocorrelation. This was done for each time point separately. Biondi et al. (1994) in considering 10-year basal area increments of pine forests in Arizona also found the empirical mark variogram decreased with increasing distance but for large distances increased. Wälder and Stoyan (1996) and Stoyan and Wälder (2000) used the same data and, focusing on short distances, concluded that the

unusual shape of the mark variogram was caused by a high frequency of pairs of dominant and suppressed trees at close proximity.

More recently developed individual-tree based models provide a powerful framework for understanding forest dynamics and spatial interaction. Suzuki et al. (2008) considered the spatio-temporal pattern of a mixed Abies forest in Japan. They used a mark variogram, a pair correlation function, and a mark correlation function with tree heights as marks for six spatial surveys over a total time period of 47 years following a large disturbance in 1959. Analysis was done for each time point separately. No human interventions occurred during this time other than the occasional removal of wind-blown trees. It is one of the few studies that show the gradual evolution of a negative autocorrelation variogram over time. The authors concluded that the variogram shape developed as a result of high mortality in clusters of suppressed trees, which increased the probability of pairs of different sized trees at close proximity.

Reed and Burkhart (1985) proposed another development path for plantation forests. In young even-aged plantations positive autocorrelation prevails. As stands develop prior to self-thinning, spatial autocorrelation becomes negative as dominant trees suppress neighboring trees, i.e. a local size hierarchy develops. Mortality of small trees as the stand experiences self-thinning eventually leads to positive autocorrelation.

Apiolaza and Garrick (2001) discuss modeling longitudinal data from measurements of weighed basic wood density of individual trees at four ages. Several models to represent the correlation matrix (unstructured, banded correlations, autoregressive, full-fit and reduced-fit random regression, repeated, and uncorrelated) are presented, and the relationships among them explained. Mehtätalo (2004) introduced a longitudinal height-diameter model for Norway spruce in Finland using a Korf growth curve and a mixed model. The number of measurement occasions for each plot of the modeling data was one to four but these covered a large age span thus making construction of the model possible. In this paper, a longitudinal model is also fitted with different correlation structures to provide a contrast to the spatial and spatio-temporal models.

The data for this study are measurements of diameter at breast height (DBH) and associated covariates of individual trees laid down in a regular grid as part of an experiment begun in 1951. These are lattice data (Cressie 1993, Section 6.3). They are distinguished from geostatistical processes by the finite size of the spatial index set D_s. However, as stated in Cressie and Wikle (2011, Section 4.5) "there are methodologies that the two spatial statistical modelling approaches (lattice and geostatistical) can share, particularly because in practice geostatistical processes are evaluated on a fine (regular) grid $s_1, \ldots, s_n$". They suggest conditional autoregressive models (CAR) for analysing lattice data but note there is always a relationship between a geostatistical model and a CAR model. We have chosen the former approach because of ease of interpretability. Similarly, Diggle and Ribeiro Jr. (2007, Section 1.2.1) also classify data such as these as discrete and suggest that pragmatically one specifies a model at the level of the discrete spatial units i.e. a multivariate distribution for random variables $Y_i : i = 1, \ldots, n$ i.e. the responses - DBH.

The objective of this paper was to develop individual tree geostatistical spatio-temporal DBH models for Sitka spruce that are dynamic and incorporate spatial structure. The dependent variable was DBH (cm) that was repeatedly measured over time. In our geostatistical model, we estimate the mean and then estimate the variogram (i.e. covariance) of the residuals (Cressie 1993, Section 2.6.2) using separate models for each. We then iterate between estimation of the mean and variogram. The variogram was estimated using weighted least squares where a parametric variogram is fitted to an empirical variogram i.e. to a binned Matheron estimator of the variogram. In contrast likelihood methods use the observed data directly and do not bin the data. Schabenberger and Gotway (2005, Section 4.5) state "No single method can claim uniform superiority." The approach has an advantage over maximum likelihood when the distribution cannot be characterized in terms of a parametric distribution function.

We fit a wave covariance function to the empirical variogram to describe the pattern of positive and negative correlation between competing trees and examine the temporal autocorrelation using data collected over a 25 year period. A spatio-temporal sum-metric

variogram was used to carry this out and using cross-validation regression kriging different variogram models are compared.

The data and models are described in Sections 2 and 3. Section 4 gives the main results and we conclude with a discussion in Section 5.

2. MATERIALS AND METHODS

2.1 Data

A thinning experiment carried out on Sitka spruce, located in a Coillte-owned (Irish Forestry Board) forest at Shillelagh, Co. Wicklow, Ireland, was modeled to examine the spatial (and spatio-temporal) dependence that may be affecting stands containing trees of the same age. The experiment was designed as a randomized complete block with three treatments representing different thinning strategies, replicated six times, resulting in a total of 18 measured plots. Planting occurred in 1951, of one year old Sitka spruce seedlings, planted with an initial espacement of 1.524 m $\times$ 1.524 m or 5 ft $\times$ 5 ft. Plots were approximately 20 m $\times$ 20 m, or one tenth of an acre, with a five metre buffer which received the same thinning treatment as the plots. The experiment was laid out in 1972, when the stand was 22 years of age. Individual-tree DBHs were recorded in each plot every 5 years from 1972 until 1997.

The three thinning treatments were unthinned, 40% thinned, and 50% thinned, with treatment beginning in 1972 and occurring on a five year cycle. The 40% thinning treatment was begun between the second and third thinning cycle (1979) and plots receiving this are referred to as 40% (delayed) plots. Initial thinning type was line thinning, and subsequent thinning type was selection. Line thinning involves the removal of a complete row of trees. In the case of 50% thinned plots every second row was removed. Selective thinning involved the forester removing trees on their relative merits, ensuring even spacing and taking trees across the range of diameters until a specified percentage of volume had been removed. Further details may be found in Lynch (1980).

The spatial coordinates of the individual trees were not recorded, but could be inferred

from the grid on which they were planted. The plots were revisited in 2012 to obtain the spatial coordinates of the remaining trees and reconstruct the original tree locations. Mortality trees contribute to estimation over the years in which they were still living and present in the data set. Their DBH was not measured after mortality occurred. Most plots contained either twelve or thirteen rows and columns. In this study, attention is restricted to one plot from each of the three treatments. This is done to illustrate the methodology in fine detail. Since plots are not located directly beside one another, spatial analyses for combined plots were not appropriate and so analyses were carried out for each plot individually and the results subsequently compared.

Individual tree DBH (cm) data for the 6 available time points were selected at the equally spaced 5-year intervals. Time points begin at 1972, when the trees were 22 years old, and end in 1997, when the trees were 47 years old. All computations were performed in R (R Core Team 2015) and models were fitted using the **gstat** package in R (Pebesma 2004).

2.2 DBH regression modeling

Three types of models were considered: (i) a regression with temporal covariance structure between repeated measures of DBH on each tree over time, (ii) a regression model with a spatial covariance structure and (iii) a regression model with a spatio-temporal covariance structure. For each model the predicted value for each data point was computed. The models were then compared using their root mean squared error (RMSE) (Section 3.4). All models were estimated using a mean based on covariates.

Dirichlet tessellation and Delaunay triangulation methods were used to calculate the area of the Dirichlet cells for each tree. Dirichlet cells were defined as follows: let $P_1, P_2, ..., P_N$ be a finite number of distinct locations in a plane. The area associated with P_N is the set T_N defined by $T_N = \{x | d(x, P_n) \leq d(x, P_m)\}$ for all $m \neq n$ where d is Euclidean distance (Okabe et al. 1992). A cell contains the ground points that are closer to that tree than to any other tree. The area of the cells or polygons (m^2) are also known as the area potentially available (APA) and are considered in a forestry context by García (2006). Trees that share

a polygon side were known as nearest neighbors. The number of nearest neighbors (NN) for each tree were computed using the **deldir** package in R (Turner 2014). Trees that were in the buffer area were included in calculating the number of nearest neighbors but were not included in variogram estimation discussed in the following sections.

A regression model was fit to individual tree DBH with the independent variables tree-age, APA and NN. Due to increasing variance of the residuals with predicted values from the model of DBH, a Box-Cox transformation of DBH was used (Box and Cox 1964). This models DBH^{λ} where λ is chosen by likelihood methods. The two different thinning types line and selective were not modeled as the selective type varied over the years and thus there was only one time-point for each selective type.

Since the data contains repeated measures on trees over time, diameter measurements from the same tree (within-subject effects) are highly correlated. In order to provide a contrast to spatio-temporal models, repeated measures analysis was carried out for the Box-Cox regression model with an unstructured covariance structure. Here, different parameters for the variance of each measurement over time were estimated as well as different covariance parameters for each pair of repeated measurements. Spatial correlation was not considered in this model.

3. SPATIAL AND SPATIO-TEMPORAL MODELS

3.1 Spatial analysis

Previous studies have identified the spherical and Gaussian variogram models as appropriate models for spatial dependence between individual trees as discussed by Fox et al. (2007*a*). To estimate spatial correlation from our data, we assumed the locations $\mathbf{s}_1, \ldots, \mathbf{s}_n$ are fixed and the observed DBH are realizations of random variables $\{Z(\mathbf{s}_1), \ldots, Z(\mathbf{s}_n)\}$ (Schabenberger and Gotway 2005). We assumed $Z(\mathbf{s})$ is composed of a mean and error

$$Z(\mathbf{s}) = m + \varepsilon(\mathbf{s}),$$

where $\mathbf{s}$ is tree location.

Diggle and Ribeiro Jr. (2007) explain that if the underlying mean function, m, is not constant, sample variograms based on the data are potentially misleading. A solution is to estimate m with a regression model. Therefore instead of a constant mean, the analysis is carried out assuming the mean m is a linear function as defined in Equation 1,

$$m = \beta_0 + \beta_1(AGE) + \beta_2(APA) + \beta_3(NN), \tag{1}$$

and then the residuals $r(\mathbf{s})$ form an isotropic stationary process.

Variability in DBH differences can be described as a function of the differences in location. This function, called the variogram, is defined as:

$$\begin{aligned} \gamma(\mathbf{s}_i - \mathbf{s}_j) &= \frac{1}{2}\mathrm{Var}\{Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\} \\ &= \frac{1}{2}\mathrm{E}\{Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\}^2. \end{aligned} \tag{2}$$

We assumed all pairs of locations $(\mathbf{s}_i - \mathbf{s}_j)$ that are a given distance $\mathbf{h}$ apart have the same variogram value, namely, $\gamma(\mathbf{h}|\boldsymbol{\theta})$, where the variogram model depends on some parameters $\boldsymbol{\theta}$, i.e. isotropic stationarity was assumed. Under this assumption, the spatial correlation of Z does not depend on location $\mathbf{s}$, but on the length of separation distance $\mathbf{h}$. Thus (2) may be written as

$$\gamma(\mathbf{h}) = \frac{1}{2}\mathrm{E}\{Z(\mathbf{s}+\mathbf{h}) - Z(\mathbf{s})\}^2.$$

An empirical variogram was computed using the Matheron (1963) estimator,

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2|N(\mathbf{h})|}\sum_{N(\mathbf{h})}\{Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\}^2, \tag{3}$$

where $|N(\mathbf{h})|$ is the number of pairs in the set $N(\mathbf{h})$ and $N(\mathbf{h}) = \{\mathbf{s}_i - \mathbf{s}_j = \mathbf{h}; i, j = 1, \ldots, n\}$. Plotting the variogram produces a curve typically rising from a point on the y-axis (the nugget), to a maximum value or 'sill' within a certain lag distance on the x-axis (the range). See Figure 1. Sample locations separated by distances closer than the range are spatially autocorrelated, whereas locations farther apart than the range are not. The nugget describes the spatially uncorrelated variation or noise in the data. The larger this value, the less spatial dependence there is amongst the attribute values.

[Figure 1 about here.]

Functions commonly used to model the sample variogram in Equation 3 include the exponential, Gaussian, spherical and the Matérn family. For these data, it was noticed especially that the largest value of $\hat{\gamma}(\mathbf{h})$ occurred at the shortest lag interval $\mathbf{h}$. The variogram then decreased and increased in a wave-like fashion. This makes sense if following the logic that, in a particular row of trees, a large (dominant) tree may be followed by a small tree. This in turn may be followed by a large tree and so on. The same applies in all directions. The wave accounts for (1) the negative spatial dependence over small inter-tree distances caused by competition among immediate neighbors (Reed and Burkhart 1985) and (2) the positive micro-site variation that occurs from trees in close proximity being exposed to the same growing conditions. If thinning has been carried out, the decreasing spatial dependence resulting from competition would be expected to diminish or become reduced over time as the trees thinned are always from the smaller diameter classes (Bradley 1971). Analysis was therefore carried out on thinned and unthinned plots so they may be compared.

Firstly a spatial analysis at each time point was carried out. This helps to determine the nature of the spatial correlation between the DBH of trees. A variety of variogram models were considered, with the wave (or *hole-effect*) model (Equation 4) performing best using the criterion sum of squares error (SSE). This is a non-monotone correlation function, they are rare in practice (Diggle and Ribeiro Jr. 2007), though the damped oscillatory nature of the function performed suitably well in this context. The wave model is given by

$$\gamma(\mathbf{h}) = \sigma_n^2 + \sigma^2(1 - (\frac{\phi}{\mathbf{h}})\sin(\frac{\mathbf{h}}{\phi})), \tag{4}$$

where σ^2 is the sill variance, σ_n^2 is the nugget variance, ϕ is the wave *intensity* and $\mathbf{h}$ is the physical distance separating two trees. The wave model is one of the few candidates for parametric models that allow for negative correlation with a small number of parameters requiring estimation. The 'practical' range for this model is defined as the lag distance at which the first peak is no greater than $1.05\sigma^2$ or the first valley is no less than $0.95\sigma^2$. It is approximately $6.5 \times \pi\phi$ (Schabenberger and Gotway 2005, page 149).

The model was fit in two stages. Firstly the regression was performed (of $Y = DBH^{\hat{\lambda}}$ on covariates) and estimates of β, i.e. $\hat{\beta}$, described in Equation 1 were obtained. Then the sample variogram was computed from the predicted residuals with all time points combined. Different variogram models were fit to this sample variogram. In most cases the wave variogram (Equation 4) provided the best fit.

3.2 Spatial prediction

Regression 2D kriging methods were used to predict values of $DBH^{\hat{\lambda}}$. If the vector of predictor values at location $\mathbf{s}_0$ are denoted by the $1 \times p$ row-vector $\mathbf{x}(\mathbf{s}_0)$, $\mathbf{V}$ is the covariance matrix of $\mathbf{Z}(\mathbf{s})$, and $\mathbf{v}$ the covariance vector of $\mathbf{Z}(\mathbf{s})$ and $Z(\mathbf{s}_0)$, then the best linear unbiased predictor of $Z(\mathbf{s}_0)$ is

$$\hat{Z}(\mathbf{s}_0) = \mathbf{x}(\mathbf{s}_0)\hat{\beta} + \mathbf{v}'\mathbf{V}^{-1}(\mathbf{Z}(\mathbf{s}) - \mathbf{X}\hat{\beta}), \tag{5}$$

where p is the number of independent variables in the regression model and $Z \equiv DBH^{\lambda}$.

Here, starting values for β, say $\hat{\beta}$, were got from ordinary least squares. The residuals, $r = Z(\mathbf{s}) - X\hat{\beta}$ were computed. An estimate of $\mathbf{V}$ was found based on on r. A new estimate of β, $\hat{\beta} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}(\mathbf{s})$, was computed (Schabenberger and Gotway 2005, Section 5.5).

If the variogram model describes adequately the spatial dependencies implicit in the dataset, then the predicted value $\hat{Z}(\mathbf{s}_0)$ should be close to the true value $Z(\mathbf{s}_0)$. Cressie (1993) discusses how, ideally, additional observations on $Z(\cdot)$ could be taken to check this, or initially some of the data might be set aside to validate the spatial predictor. Here however, all of the data were used to fit the variogram and to build the spatial predictor, and there was no possibility of taking more observations. A cross-validation approach was used to assess model fit.

Leave-one-out cross-validation (LOOCV) of the kriging predictions was used to compute the RMSE for these models. This provides unbiased estimates of $Z(\mathbf{s}_i)$. Let $\hat{\gamma}(\mathbf{h})$ be the fitted variogram model (obtained from all the data). Now delete a datum $Z(\mathbf{s}_i)$ and predict it with $\hat{Z}(\mathbf{s}_i)$ using Equation 5 based on $\hat{\gamma}(\mathbf{h})$ and the data $Z_{(-i)} = (Z_1, ..., Z_{i-1}, Z_{i+1}, ..., Z_n)$.

This is done for $i = 1, ..., n$, where n is the number of observations in a plot. The prediction error can be inferred from the predicted-minus-actual values, i.e. the RMSE (Chilès and Delfiner 1999, page 290),

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}[\hat{Z}(\mathbf{s}_j) - Z(\mathbf{s}_j)]^2}, \tag{6}$$

where $\hat{Z}$ is on the back-transformed scale (cm). If the model fits well, cross-validation residuals should be small, have zero mean, and no apparent structure. It should be noted that in the cross-validation procedure, the variogram model is not re-fit for each leave-one-out fit. A variogram is fit on the complete data set, and in that case validation residuals are not completely independent from modeling data, as they already did contribute to the variogram model fitting. Note model fit statistics cannot be compared using criteria such as AIC as we did not specify a likelihood and use maximum likelihood. Instead we compare models using RMSE as in Bivand, Pebesma and Gómez-Rubio (2008, Section 8.6).

3.3 Spatio-temporal analysis

Joint analysis of spatio-temporal data are preferable to separate analysis. Interpolation of observations in a continuous spatio-temporal process should take into account the interactions between the spatial and temporal components and allow for predictions in time and space. Separate analysis of these processes allow predictions in space or time only. Kyriakidis and Journel (1999) describe how 2-dimensional spatial models can be extended to the time domain leading to spatio-temporal geostatistics. However, if time is considered an 'added' dimension, then anisotropy between space and time must be taken into account. The space-time variation of DBH can be characterized by decomposing it into a deterministic trend m and a zero-mean stochastic error term ε as follows:

$$Z(\mathbf{s}, t) = m(\mathbf{s}, t) + \varepsilon(\mathbf{s}, t).$$

The trend m is the same as Equation 1. We again assume that the zero-mean stochastic residual term ε is multivariate normally distributed. Given this assumption, the only

information lacking is its autocovariance function,

$$C(\mathbf{s}_i, t_i, \mathbf{s}_j, t_j) = \mathrm{E}[\varepsilon(\mathbf{s}_i, t_i) \times \varepsilon(\mathbf{s}_j, t_j)]. \tag{7}$$

The specification (7) cannot be estimated because, in practice, we have only one realization from which to infer a covariance function. To facilitate the estimation of C from observations, some limiting additional simplifying assumptions are necessary. Therefore, as previously discussed, a stationarity assumption is made, which posits that the covariance of $\varepsilon(\mathbf{s}, t)$ and $\varepsilon(\mathbf{s} + \mathbf{h}, t + u)$ depends only on the distance in space $\mathbf{h}$ and distance in time u between the points: $C(\mathbf{s}, t, \mathbf{s} + \mathbf{h}, t + u) = C(\mathbf{h}, u)$.

One major difficulty is to ensure that the space-time covariance function is valid (i.e., positive definite). Here we considered a positive-definite nonseparable space-time covariance function with a model-based approach. We defined a model for ε and derived the associated covariance function from it. If each of the spatial, temporal and joint spatio-temporal components of the model has a positive-definite covariance function, the positive definiteness of the space-time covariance function is guaranteed (Bilonick 1988). We use:

$$\varepsilon(\mathbf{s}, t) = \varepsilon_s(\mathbf{s}) + \varepsilon_t(t) + \varepsilon_{st}(\mathbf{s}, t),$$

where ε_s is a purely spatial process (i.e., its realizations are constant over time), ε_t is a purely temporal process (i.e., realizations are constant in space), and ε_{st} is a space-time process for which distance in space is made comparable to distance in time by introducing a space-time anisotropy ratio. All three components of the above equation are assumed stationary and mutually independent, which leads to the 'sum-metric' space-time covariance structure

$$C(\mathbf{h}, u) = C_s(\mathbf{h}) + C_t(u) + C_{st}(\sqrt{\mathbf{h}^2 + (\alpha \times u)^2}), \tag{8}$$

where α is an anisotropy ratio introduced to make distance in space comparable to distance in time. The first two terms in the right-hand side of Equation 8 allow for the presence of zonal anisotropies (i.e., variogram sills that are not the same in all directions). Zonal anisotropy occurs when the amount of variation in time is smaller or greater than that in

space and/or that in joint space-time. The geometric anisotropy ratio α that appears in the third term in the right-hand side of Equation 8 is needed because a unit of distance in space is not the same as a unit of distance in time. For instance, if α =20 m per day, then two points that are separated by 100 m in space and zero days in time have the same correlation as two points that are five days apart in time and zero meters in space, or as two points that are separated by 60 m in space and four days in time (Heuvelink and Griffith 2010, page 167).

The sample spatio-temporal variogram of the residuals from the Box-Cox regression model was estimated for each plot. The sum-metric covariance function was used to fit a spatio-temporal variogram model to each sample variogram. Different choices for C_s, C_t and C_{st} were considered. The sum-metric covariance model was fitted using the `fit.StVariogram` function available in the **gstat** package in R. We first determined initial estimates for: (1) the nugget of the marginal temporal variogram, (2) sill of marginal temporal variogram, (3) temporal range parameter, (4) nugget of marginal spatial variogram, (5) sill of marginal spatial variogram, (6) spatial range parameter and (7) total sill, by examining plots of the marginal sample variograms. Parameters were then determined using the `optim` function in R until a satisfactory fit was achieved. The procedure to estimate a sum-metric spatio-temporal variogram is explained in more detail in Heuvelink and Griffith (2010). The fitted variogram model (i.e. estimated covariance matrix $\mathbf{V}$) was used to carry out spatio-temporal regression kriging.

The model which performed best, for each plot, consisted of a wave function for the spatial process, the temporal process and the spatio-temporal process (Equation 9),

$$\begin{aligned}\gamma(\mathbf{h}, u) = {} & \sigma_{n1}^2 + \sigma_1^2[1 - ((\frac{\phi_1}{\mathbf{h}})\sin(\frac{\mathbf{h}}{\phi_1}))] \\ & + \sigma_{n2}^2 + \sigma_2^2[1 - ((\frac{\phi_2}{u})\sin(\frac{u}{\phi_2}))] \\ & + \sigma_{n3}^2 + \sigma_3^2[1 - ((\frac{\phi_3}{\sqrt{\mathbf{h}^2 + (\alpha \times u)^2}})\sin(\frac{\sqrt{\mathbf{h}^2 + (\alpha \times u)^2}}{\phi_3}))],\end{aligned} \tag{9}$$

where σ_{n1}^2, σ_{n2}^2 and σ_{n3}^2 are the spatial, temporal and joint spatial-temporal nuggets, respec-

tively, σ_1^2, σ_2^2 and σ_3^2 are the respective partial sills, ϕ_1, ϕ_2 and ϕ_3 are the wave intensity parameters and α is the anisotropy ratio. The sum-metric model was compared to the sum model, which excludes the third term of Equations 8 and 9, using the criterion sum of squares error (SSE) and was found to fit better. This model differs from other space-time models by the inclusion of a space-time interaction.

3.4 Spatio-temporal prediction

The kriging Equation 5 is the same for spatial and spatio-temporal data. DBH predictions were made using regression kriging (Box-Cox model) with the fitted spatio-temporal variogram models described in Equation 9.

Similarly to the spatial prediction, leave-one-out cross-validation of the kriging predictions was carried out as a means of comparing our model choices. For example, for an unthinned plot there are 6 time points for each tree, i. Each tree i was deleted in turn (i.e. all of its 6 time points) and Equation 5 used to predict ($\hat{Z}(\mathbf{s}_i, t_j)$, $j = 1, \ldots, 6$), where the matrix $\mathbf{V}$ used in the prediction is based on all other trees and all their 6 time points. This method follows Hengl et al. (2012). It can be justified in the sense that if no observation is taken at a location $\mathbf{s}_i$, say, then one uses all available information to predict it. However, LOOCV was also carried out by a second method based on predicting a tree at time j, while including that tree's values at times previous to time j, together with the other trees.

The spatio-temporal variogram modeling and kriging was implemented in R using the **gstat** package. We compared the results of spatial with spatio-temporal cross-validation kriging predictions using root mean squared prediction error (Equation 6).

4. RESULTS

Trends in diameter through the six time points for different thinnings are shown in Figure 2. The distribution of DBH in 1972 indicates, as expected, no difference in DBH range at the start of the experiment for the three treatments. The distribution of values over time indicates that higher levels of thinning resulted in improved DBH growth in retained trees. The no thinning treatment is associated with the lowest mean DBH because dead and

suppressed trees are retained with the subdominant and dominant trees. Higher intensities of thinning are associated with an increase in the mean DBH because trees were removed during thinning generating advanced DBH growth in retained trees. The distribution of DBH at the six time points, from 1972 to 1997, clearly illustrate the very positive effect of both thinning treatments on the diameter distribution.

[Figure 2 about here.]

Figure 3 presents the spatial coordinates of individual tree locations for the unthinned plot, at time points 1972 and 1997, and for the 50% thinned plot, at time point 1997. The unit of measurement for both the x and y coordinates is the metre. The circle radii are proportional to the DBH (cm). The Dirichlet tessellation is illustrated by solid lines and defines the polygon and the area potentially available (APA) (m^2) for each tree. The number of sides of a polygon represents the number of nearest neighbors (NN) for the tree contained within it.

[Figure 3 about here.]

At the first time point in 1972 the unthinned plot contained 169 trees. This was reduced to 77 trees by 1997. An average of 15 trees died every 5 years due to natural mortality from factors such as disease and a shorter life-span of suppressed trees. Some trees also suffered windblow damage leading to mortality.

4.1 DBH regression model with repeated measures

The Box-Cox regression model, Section 2.2, was fit to the data from the unthinned plot. The mean, m, was estimated using a linear model with predictor variables: AGE, APA and NN, Equation 1. Initially a Box-Cox transformation was used, $Y = DBH^{\hat{\lambda}}$. The estimate of λ was 0.384, which was chosen by likelihood methods. The effect of the transformation on the residuals can be seen in Figure 4. The estimated coefficients for AGE, APA and NN were all positive and very highly signifcant ($P < 0.001$). The RMSE for the regression

model is 4.128. After repeated measures analysis was carried out the RMSE value increased to 5.774 (Table 1).

[Table 1 about here.]

[Figure 4 about here.]

The thinned plots (40% and 50%) contained 163 and 166 tree stems in 1972. This was reduced to 24 and 19 trees, respectively, by 1997. The percentage of trees lost to mortality (15% and 0% respectively) was considerably lower than in the unthinned case (55%). A Box-Cox regression model was also fitted to each of these plots and the values of $\hat{\lambda}$ were 0.585 and 0.747, respectively. The RMSE values for the regression model with repeated measures are also given in Table 1. In all three models, the estimated coefficients for AGE and APA were significant ($P < 0.001$), while the coefficients for NN were significant for the unthinned plot and the plot which received a 50% thinning, but not for the plot which received a 40% thinning. The standard errors for the estimated coefficients take into account correlation in the data over time although not over space. This is done later in a more complete and parsimonious way using spatial and spatio-temporal covariance models. The regression model for the mean used here for DBH had R^2 values of 0.44, 0.67 and 0.85 for the three (unthinned, 40% thinned and 50% thinned) respective plots.

4.2 DBH spatial analysis

We examined the spatial covariance of the transformed DBH regression residuals for each of the three plots. The sample variogram for the residuals based on the transformed model was computed with all time points combined. A variety of covariance functions were fit to the spatial variograms with the wave model giving the best fit (SSE comparison) in each case. The variogram model parameters (Table 2) show a sill of 0.08 and a substantial nugget value of 0.053 in the unthinned plot. This is expected to be the case as the variance at spatial lag zero includes repeated diameter measurements from the same tree taken at each time point while any temporal effect is ignored except that built into the mean structure.

A similar result is seen in the 40% (delayed) and 50% thinned plots. Figure 5 illustrates the fitted spatial variogram for the 50% thinned plot. The spatial range parameter values are small at 1.7 m, 5.8 m, and 3.7 m, for the unthinned, 40% thinned, and 50% thinned plots, respectively, showing that the residuals are only correlated with their closest neighbors. The variogram model was used to carry out 2D, spatial, leave-one-out cross-validation (LOOCV) regression kriging as described in Section 3.2. We carried out regression kriging in an attempt to remove temporal (age) and spatial (APA and NN) trends in the data that might account for the spatial (or spatio-temporal, see below) behavior of the DBH covariances. The RMSE for the predicted values are 4.206 (unthinned plot), 4.266 (40% thinned plot), and 3.969 (50% thinned plot), as shown in Table 1). The predicted values were back-transformed before calculating the RMSE.

For the thinned plots, thinning did not take place until the first time point in 1972. The thinning effect was not yet manifest in the diameter data and the variogram model parameters essentially provided a description of an unthinned treatment. For this reason, 1972 data from thinned plots were not included in any analysis. The spatio-temporal variogram contains 5 time points which start in 1977 and end in 1997.

[Table 2 about here.]

[Figure 5 about here.]

4.3 DBH spatio-temporal analysis

The spatio-temporal sample variograms for the transformed DBH residuals are presented in Figures 6 and 7 for each of the three plots. After exhaustively checking for appropriate initial values, the sum-metric model best described the spatio-temporal covariance with a wave function used for the spatial component, temporal component and joint spatio-temporal component. The wave form over time is induced by the thinning selection strategy or by natural selection in the case of the unthinned plot. If a tree has small DBH at time t_0, it will have a relatively larger DBH at time t_1 due to thinning of trees near it. Similarly, the wave pattern over space is induced by thinning selection strategy or natural selection. In the

variogram model, a wave component for time fitted better than just a nugget component for time or any other parametric variogram form (using SSE criterion).

The model parameters are shown in Table 3. For the unthinned plot, the spatial sill (0.056) is larger than the temporal sill (0.006), which confirms that spatial variation in the transformed DBH residual dominates temporal variation. The nugget variance for the temporal component is zero, which implies that the DBH residuals are perfectly correlated at very short distances in time. Yet the space-time component has a sill larger than the temporal sill (0.047), and its contribution needs to be considered in assessing temporal correlation.

Similar to the spatial variogram models in Section 4.2, the range of the spatial component in the unthinned plot is also a short distance of 1.24 m, however this increases when the joint spatio-temporal range is accounted for. The temporal range is 15 showing that residuals are correlated up to a time period of 15 years.

There is little anisotropy in the unthinned case (0.02 m/yr). Since the anisotropy parameter, α, is multiplied by the temporal lag, u (see Equation 8), this suggests the joint spatio-temporal component contributes more to the spatial aspect than the temporal aspect and the sum-metric variogram model essentially becomes the sum model. Here, the differences in the covariances through space are the same over time and the shape of the variogram does not change with increasing temporal lags.

The anisotropy parameter, α, provides a unique estimate of the spatio-temporal homogeneity of the DBH covariance structure in thinned and unthinned stands. The anisotropy parameter, α, is much larger for the thinned plots, 0.19 and 0.18 m/yr, by a factor of approximately nine times compared to the unthinned anisotropy parameter, α, 0.02 m/yr, though there is very little difference between the 40% and 50% case. In the 50% thinned case, it means, for example, that two trees measured at the same point in time and separated by 10 meters have, on average, the same space-time correlation (partial sill) as a single tree separated by about 55.6 years in age, or two trees separated by 5 meters and 19.35 years.

As in the spatial case, the variogram models fit to the residuals from the Box-Cox

regression model were used to carry out 3D, spatio-temporal, leave-one-out cross-validation regression kriging as described in Section 3.3.

[Table 3 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

4.4 Model comparison

The three fitted models, for each plot, were compared using the RMSE in Table 1 as discussed in Section 2.2). For the unthinned plot, the root mean square (predictive) error was reduced from 5.774 to 4.206 after spatial regression kriging was carried out. Similarly, for spatio-temporal variogram modeling, regression kriging performed better than the repeated measures regression model, however, the RMSE value is almost the same as that of the spatial (2D) models. The results show that the average accuracy for predicting tree diameters is ± 4.2 cm. Similarly, for the 40% and 50% thinned plots the RMSE was reduced going from a repeated measures model to a spatial or spatio-temporal model. Again there was little difference in the fit of the spatial and spatio-temporal models. If, however, the RMSE in the spatio-temporal model is calculated using the second method of Section 3.4 then values achieved are much smaller than the other models. The unthinned plot had an RMSE of 2.426, the 40% thinned plot had an RMSE of 3.087 and the 50% thinned plot had an RMSE of 2.924. The spatio-temporal model did better much better then as it takes into account correlation in tree values over time as well as space. This assumes previous measurements on a tree are available, unlike the first method that assumes predicting at a location where no measurements are available.

The RMSE values for the spatial and spatio-temporal kriging predictions were very close, however, it must be remembered that repeated measurements were present but ignored in the spatial model. The measurements are correlated as they are from the same tree and thus the RMSE in the spatial model is an underestimate. Omitting the last term in Equation 5,

prediction was limited to the trend component, $\mathbf{X}\hat{\beta}$, ignoring the prediction of the residual i.e. generalized least squares (GLS) trend estimation was done. Values were then back transformed and the average bias for GLS, for example, for the unthinned plot, was 2.871 cm (RMSE = 5.199 cm). Comparing this to the spatio-temporal model with a bias of 1.433 cm (RMSE = 4.205 cm), spatio-temporal kriging improved bias and RMSE.

The observed, spatial predicted, and spatio-temporal predicted cumulative basal area (CBA) (m^2) per hectare after thinning, inclusive of all previous thinnings, are presented in Table 4 for each of the three plots.

[Table 4 about here.]

For the no thinning treatment, the predicted CBA from regression kriging underestimated the observed CBA. The spatial regression kriging predicted CBA was closer to the observed than that from the spatio-temporal predicted CBA. For the thinned treatments, the spatial regression kriging led to an over-prediction of CBA, while the spatio-temporal predicted CBA was closer to the observed value for the 40% thinned treatment and only slightly closer than the spatial case for the 50% thinned treatment. The observed, spatial predicted, and spatio-temporal predicted cumulative basal area (CBA) per hectare (m^2ha^{-1}) after thinning, for age from 22 to 47 years, are presented in Figure 8 for each of the three plots. CBA here refers to the total across all trees' predicted DBHs at a given point in time with also the predicted DBHs of trees that were thinned at previous time points (the thinned plots only). Thus CBA at a point in time represents 'timber volume' from the plot to date. Mortality trees are not included as they don't contribute to timber volume. CBA declines in the spatio-temporal case due to trees lost to mortality.

[Figure 8 about here.]

5. DISCUSSION

The main objective of spatio-temporal modeling was not only to predict DBH but quantify individual tree yield and inter-tree dependence. It was not possible to incorporate biological

growth and yield functions, such as the Chapman-Richards (CR) function, into the mean regression model as data was only available for the trees from age 22 onwards. It is known that the second derivative of the CR function, with respect to age, defines the juvenile, adolescent, mature and senescent stages of growth start and end and these are useful quantities in comparing thinning treatments. Future work will focus on extending CR models to the spatio-temporal domain.

It was seen in both spatial and spatio-temporal models that the variograms for all three plots show a pronounced wave effect due to competition between trees as described earlier. The wave had a higher frequency in the unthinned plot (Figure 7) where there is greater natural competition.

The spatio-temporal model fit to the 50% thinned plot was less accurate for large time lags and this may have resulted in over prediction of the total basal area as seen in Table (4). However, in general, the model fitted well, as can be seen in terms of the RMSE (Table 1) and predicted total BA (Table 4) values.

The spatial models performed almost as well as the spatio-temporal models indicating spatial variability is more important in these plots than temporal variability. This can be seen in Table (3) where temporal sills are small relative to spatial sills.

This study quantified spatial variation within 0.04 ha permanent sample plots which were subject to three thinning treatments over a 25 year period. In terms of fixed effects, it was assumed there was a constant-in-time thinning effect which was incorporated in the covariance structure. It is possible that quantifying the thinning effect spatially and temporally and including it as a fixed effect may yield similar fits as the spatio-temporal models. Watson (1972) and Ford and Diggle (1981) state it is not always possible to distinguish between models with a spatial trend and those with a spatial covariance structure. The advantage of using a covariance modelling approach here allowed the identification of the space-time interaction.

In terms of fitting the regression kriging models, a fixed effects model was first fitted and then a variogram model fitted to the residuals. This is the first iteration in simultaneously

fitting a mean and covariance structure but Schabenberger and Gotway (2005, Section 6.3) indicate that this will give a close approximation to a fully iterated model. O'Rourke (2015) found a similar result.

The mean structure was to some extent limited by the scope of the data. While for some trees (volume-sample tree) data containing the sectional lengths and diameters were available, including repeated measurements, these data were not available for all trees in the plot. It was essential from the spatial viewpoint to include all trees in the analysis so that 'competition' between adjacent trees could be modeled. Modeling DBH increment or including lagged DBH values in the mean model was not done for several reasons. The spatio-temporal variogram estimator based on DBH increments is complex to interpret. The spatio-temporal variogram based on actual DBH is already based on differences of DBH over time (i.e. increment) and differences in space. In addition, the number of available time-points for each tree is limited to six. No previous measurements are available for the first time-point so this would have to be excluded from a model with lagged DBH reducing the size of the data set. There is also the question of what is the right lag *structure*. Not only might a previous measurement of DBH be a good predictor but two or three lagged measurements might be. Again these data are not available for all time-points. Moreover, including lagged DBH in the model induces a complex autocorrelation into the residuals over time that is difficult to interpret in terms of variogram modeling.

Another problem that confronts the modeling is that observations are generally spread unevenly over space and time. The lack of temporal data may change in future with high resolution satellite or areal imagery or LIDAR remote sensing technology (Dubayah and Drake 2000) facilitating measurement.

The cited RMSE values for the spatial and spatio-temporal models, while better than the repeated measures model, are still quite high. Including more extensive data such as tree heights or possibly more time points might have reduced these. Also, a more detailed mean model, such as the Chapman-Richards function discussed above, may lead to improved RMSE values. However, we note the RMSE values for the spatio-temporal models are

considerably reduced if prediction includes previous measurements on a tree (not just on other trees).

This is a novel study in terms of modeling forestry data and represents to the best of our knowledge, the first use of spatio-temporal models in this context. The models are useful in that they illustrate the competition over space and time. It is clear that spatio-temporal dependence has an important impact on tree growth and models need to reflect this.

ACKNOWLEDGEMENTS

The authors would like to thank an Associate Editor and three referees whose careful reading and comments improved this manuscript.

REFERENCES

Apiolaza, L. A., and Garrick, D. J. (2001), "Analysis of longitudinal data from progeny tests: some multivariate approaches," *Forest science*, 47(2), 129–140.

Bilonick, R. A. (1988), "Monthly hydrogen ion deposition maps for the northeastern U.S. from July 1982 to September 1984 ," *Atmospheric Environment (1967)*, 22(9), 1909–1924.

Biondi, F., Myers, D. E., and Avery, C. C. (1994), "Geostatistically modeling stem size and increment in an old-growth forest," *Canadian Journal of Forest Research*, 24(7), 1354–1368.

Bivand, R. S., Pebesma, E., and Gómez-Rubio, V. (2008), *Applied Spatial Data Analysis with R*, Use R!, New York: Springer.

Box, G., and Cox, D. (1964), "An analysis of transformations," *Journal of the Royal Statistical Society*, 26(2), 211–252.

Bradley, R. (1971), *Thinning Control in British Woodlands* Forestry Commission Booklet No. 32.

589 Broad, L. R., and Lynch, T. (2006), "Growth models for Sitka spruce in Ireland," *Irish*
590 *Forestry*, 63, 53–79.

591 Chilès, J. P., and Delfiner, P. (1999), *Geostatistics: Modeling Spatial Uncertainty, Second*
592 *Edition* Wiley, New York.

593 Cressie, N. A. C. (1993), *Statistics for Spatial Data (Revised Edition)* Wiley-Interscience.

594 Cressie, N. A. C., and Wikle, C. K. (2011), *Statistics for Spatio-temporal Data* John Wiley
595 & Sons.

596 Diggle, P. J., and Ribeiro Jr., P. J. (2007), *Model-based Geostatistics* Springer.

597 Dubayah, R. O., and Drake, J. B. (2000), "Lidar remote sensing for forestry," *Journal of*
598 *Forestry*, 98(6), 44–46.

599 Ford, E. D., and Diggle, P. J. (1981), "Competition for Light in a Plant Monoculture
600 Modelled as a Spatial Stochastic Process," *Annals of Botany*, 48(4), 481–500.

601 Forest Service (2007), *National Forest Inventory of Republic of Ireland* Forest Service, De-
602 partment ofAgriculture, Fisheries and Food.

603 Fox, J. C., Bi, H., and Ades, P. K. (2007*a*), "Spatial dependence and individual-tree
604 growth models: I. Characterising spatial dependence," *Forest Ecology and Manage-*
605 *ment*, 245(13), 10–19.

606 Fox, J. C., Bi, H., and Ades, P. K. (2007*b*), "Spatial dependence and individual-tree
607 growth models: II. Modelling spatial dependence," *Forest Ecology and Management*,
608 245(13), 20 – 30.

609 García, O. (2006), "Scale and spatial structure effects on tree size distributions: implications
610 for growth and yield modelling," *Canadian Journal of Forest Research*, 36(11), 2983–
611 2993.

Hengl, T., Heuvelink, G., Perčec Tadić, M., and Pebesma, E. (2012), "Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images," *Theoretical and Applied Climatology*, 107(1-2), 265–277.

Heuvelink, G. B. M., and Griffith, D. A. (2010), "Space-Time Geostatistics for Geography: A Case Study of Radiation Monitoring Across Parts of Germany," *Geographical Analysis*, 42(2), 161–179.

Keddy, P. A. (1989), *Competition* Chapman and Hall, London.

Kyriakidis, P., and Journel, A. (1999), "Geostatistical Space-Time Models: A Review," *Mathematical Geology*, 31(6), 651–684.

Lynch, T. J. (1980), "Thinning and spacing research in Sitka spruce and Lodgepole pine," *Irish Forestry*, 37(2), 45–67.

Matern, B. (1960), "Spatial variation. Stochastic models and their application to some problems in forest surveys and other sampling investigations.," *Meddelanden fran Statens Skogsforskningsinstitut*, 49(5), 144.

Matheron, G. (1963), "Principles of geostatistics," *Economic Geology*, 58(8), 1246–1266.

Matthews, R., Mackie, E., and Hamilton, G. J. (2006), *Forest Mensuration: A handbook for practitioners* Forestry Commission.

Mehtätalo, L. (2004), "A longitudinal height diameter model for Norway spruce in Finland," *Canadian Journal of Forest Research*, 34(1), 131–140.

Okabe, A., Boots, B., and Sugihara, K. (1992), *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, New York, NY, USA: John Wiley & Sons, Inc.

O'Rourke, S. (2015), Spatial and Spatio-temporal Modelling of Sitka spruce Tree Growth from Forest Plots in Co. Wicklow, PhD thesis, University College Dublin.

Pebesma, E. J. (2004), "Multivariable geostatistics in S: the gstat package," *Computers & Geosciences*, 30, 683–691.

Pommerening, A., and Särkkä, A. (2013), "What mark variograms tell about spatial plant interactions," *Ecological Modelling*, 251, 64–72.

R Core Team (2015), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Reed, D. D., and Burkhart, H. E. (1985), "Spatial Autocorrelation of Individual Tree Characteristics in Loblolly Pine Stands," *Forest Science*, 31(3), 575–587.

Schabenberger, O., and Gotway, C. A. (2005), *Statistical Methods for Spatial Data Analysis* Chapman & Hall/CRC Press.

Stoyan, D., and Wälder, O. (2000), "On Variograms in Point Process Statistics, II: Models of Markings and Ecological Interpretation," *Biometrical Journal*, 42(2), 171–187.

Suzuki, S. N., Kachi, N., and Suzuki, J.-I. (2008), "Development of a Local Size Hierarchy Causes Regular Spacing of Trees in an Even-aged Abies Forest: Analyses Using Spatial Autocorrelation and the Mark Correlation Function," *Annals of Botany*, 102(3), 435–441.

Turner, R. (2014), *deldir: Delaunay Triangulation and Dirichlet (Voronoi) Tessellation.* R package version 0.1-7.

Wälder, O., and Stoyan, D. (1996), "On Variograms in Point Process Statistics," *Biometrical Journal*, 38(8), 895–905.

Watson, G. S. (1972), "Trend surface analysis and spatial correlation," *Geological Society of America Special Papers*, 146, 39–46.

List of Figures

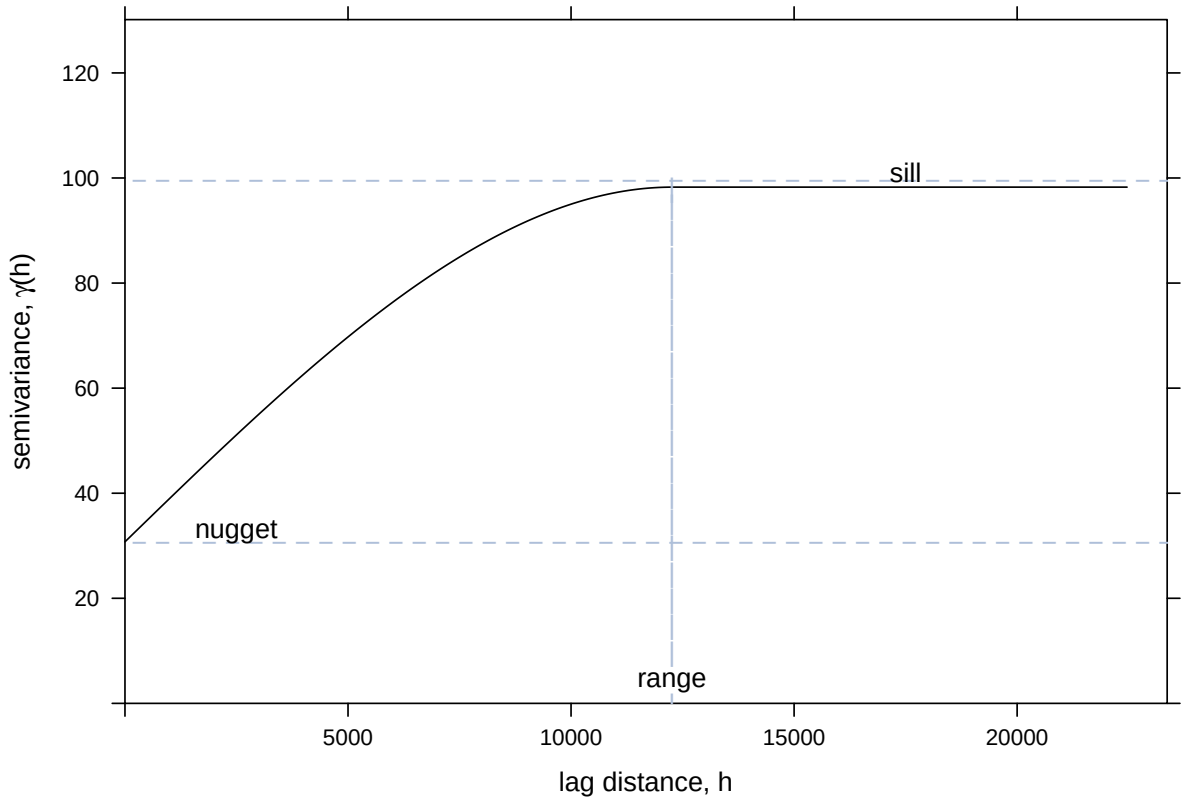


Figure 1: Schematic representation of a typical variogram, with structural parameters indicated.

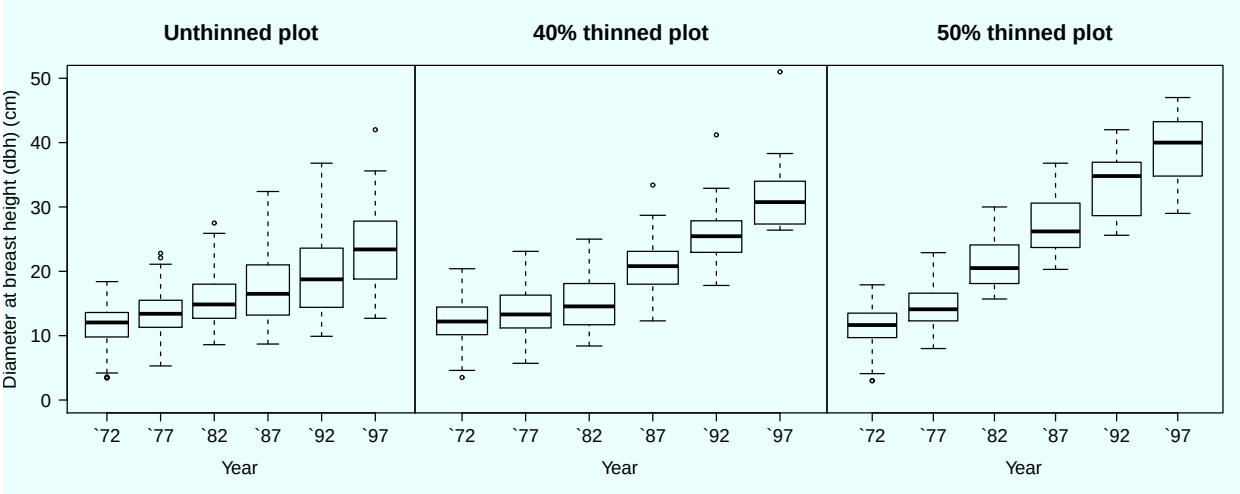
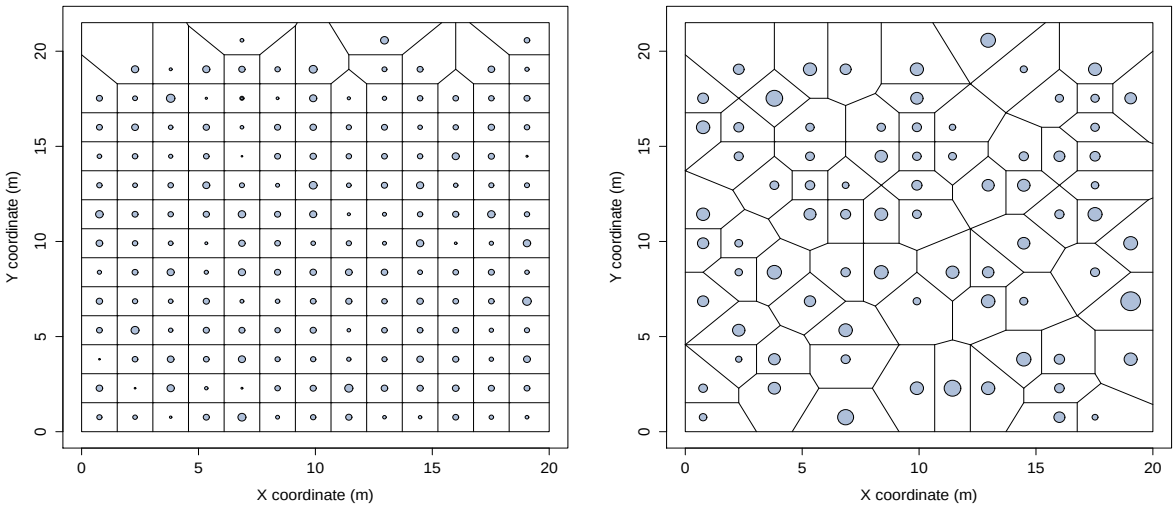


Figure 2: Box plots for tree diameters at each time point.

(a) Unthinned plot in 1972

(b) Unthinned plot in 1997



(c) 40% thinned plot in 1997

(d) 50% thinned plot in 1997

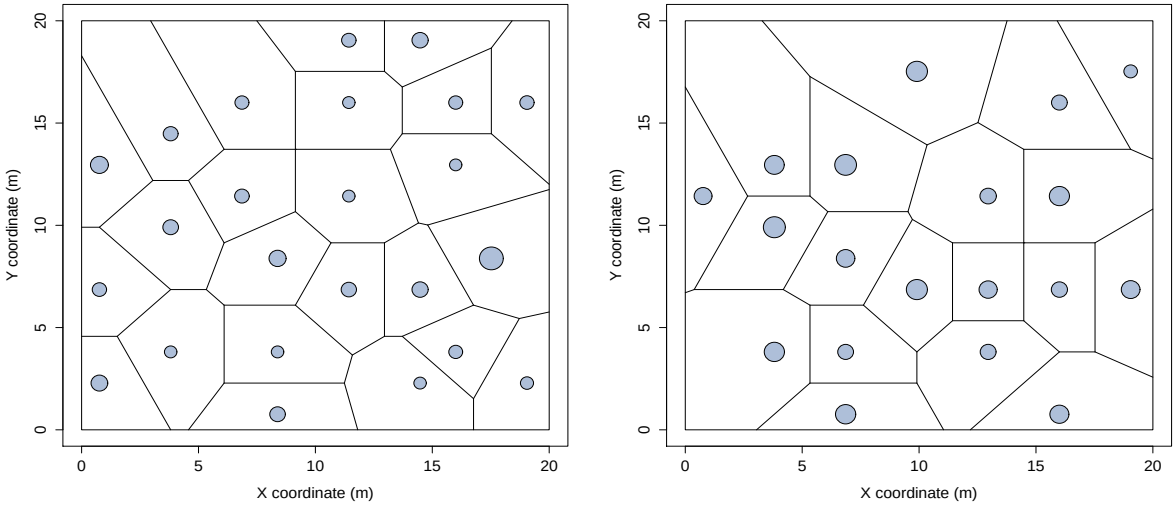
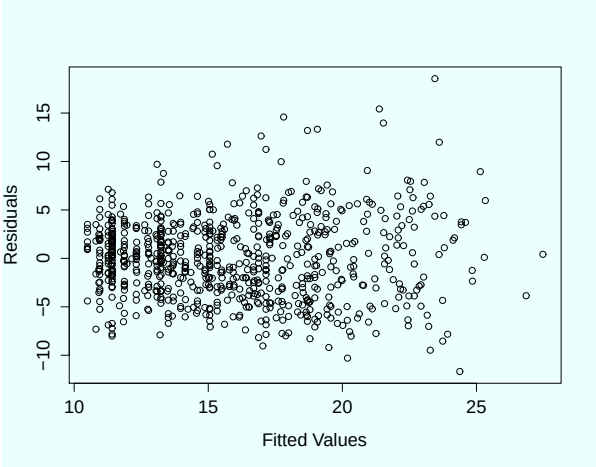
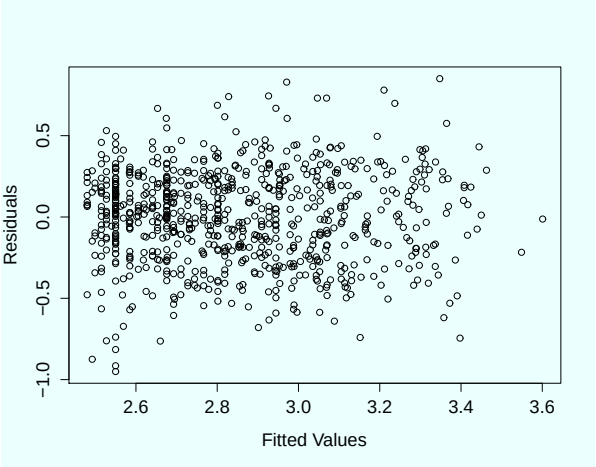


Figure 3: Circle plots showing tree locations where circle radii are determined by DBH. Tessellation lines determine polygons with the number of sides representing the number of nearest neighbors (NN) for the tree contained within it.



(a) Unthinned plot, Y=DBH (cm)



(b) Unthinned plot, Y=DBH$^{\hat{\lambda}}$ (cm$^{\hat{\lambda}}$)

Figure 4: Regression model residuals (cm) vs. fitted values (cm) for the unthinned plot (a) without Box-Cox transformation and (b) with Box-Cox transformation (cm$^{\hat{\lambda}}$).

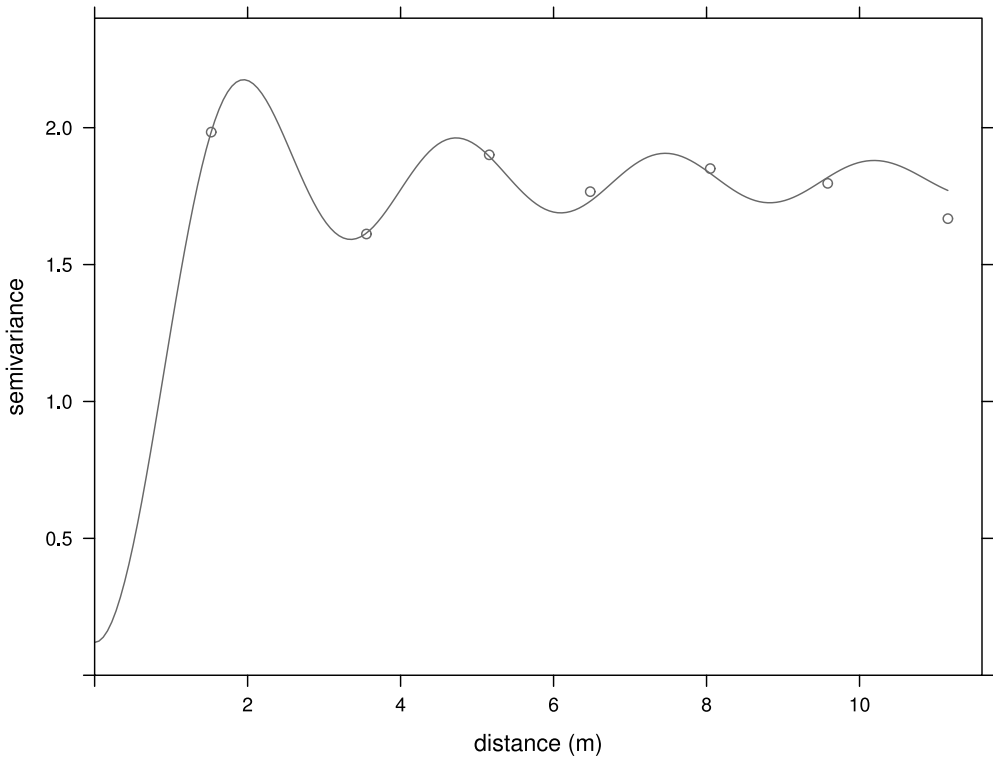
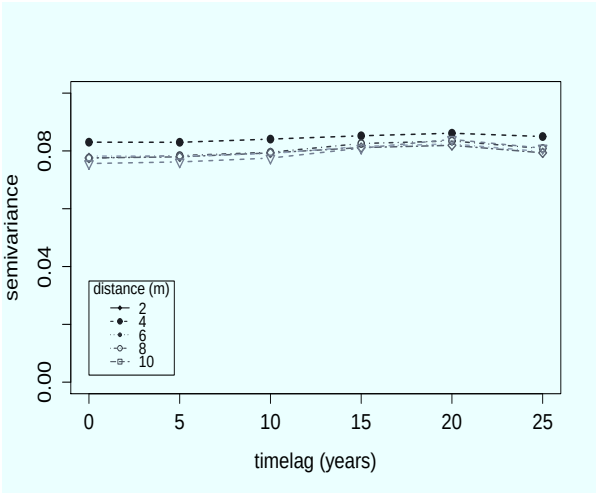
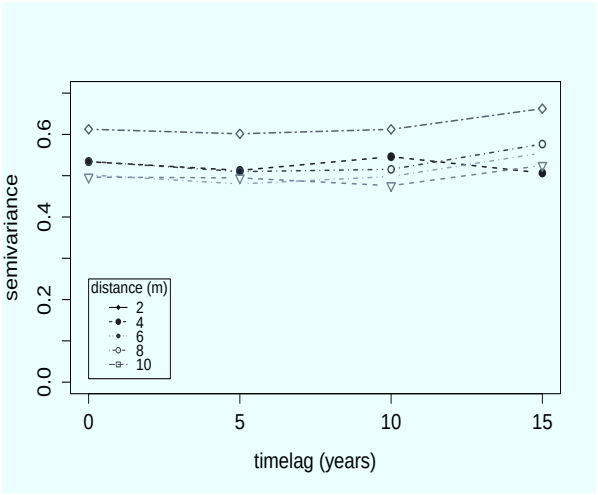


Figure 5: Variogram of transformed regression residuals (dots) with fitted wave model (line) for a 50% thinned plot for all time points combined.

(a) (b)





(c)

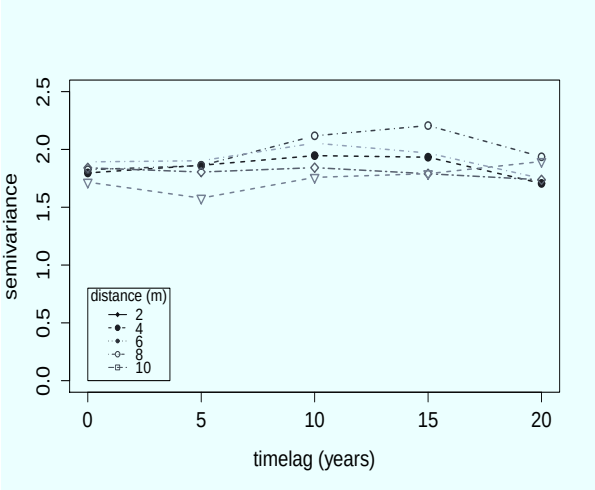
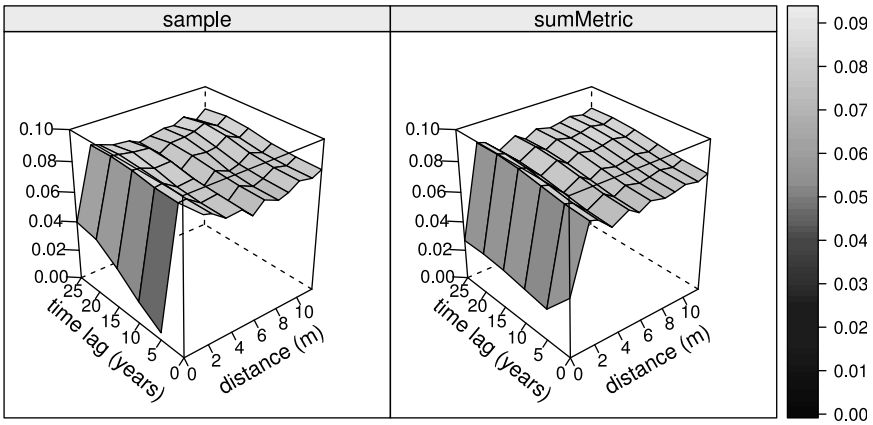
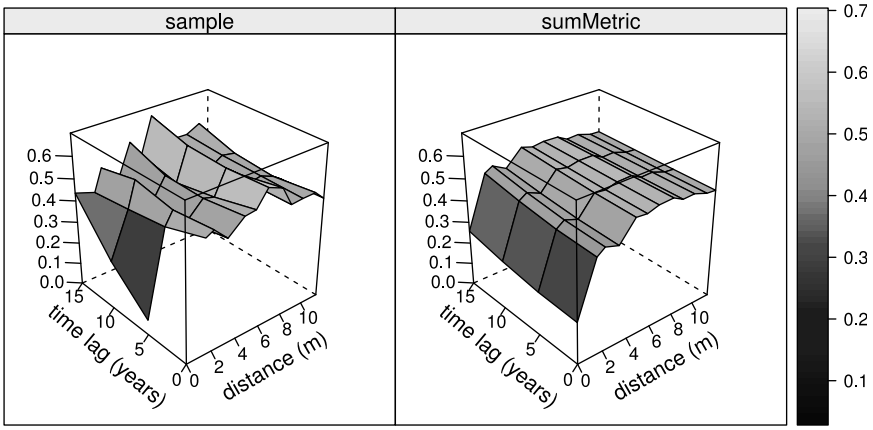


Figure 6: Spatio-temporal sample semivariogram for residuals from the regression model at different distances in (a) an unthinned plot, (b) a 40% thinned plot, and (c) a 50% thinned plot, with time lag on the x-axis. The wave pattern over time can be seen.

(a)



(b)



(c)

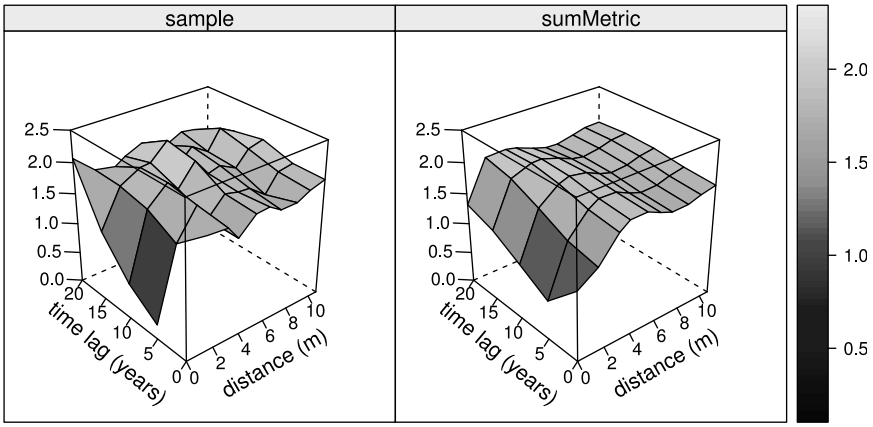


Figure 7: 3D Spatio-temporal sample variogram and sum-metric variogram model for transformed residuals of tree diameters in (a) an unthinned plot, (b) a 40% thinned plot and (c) a 50% thinned plot.

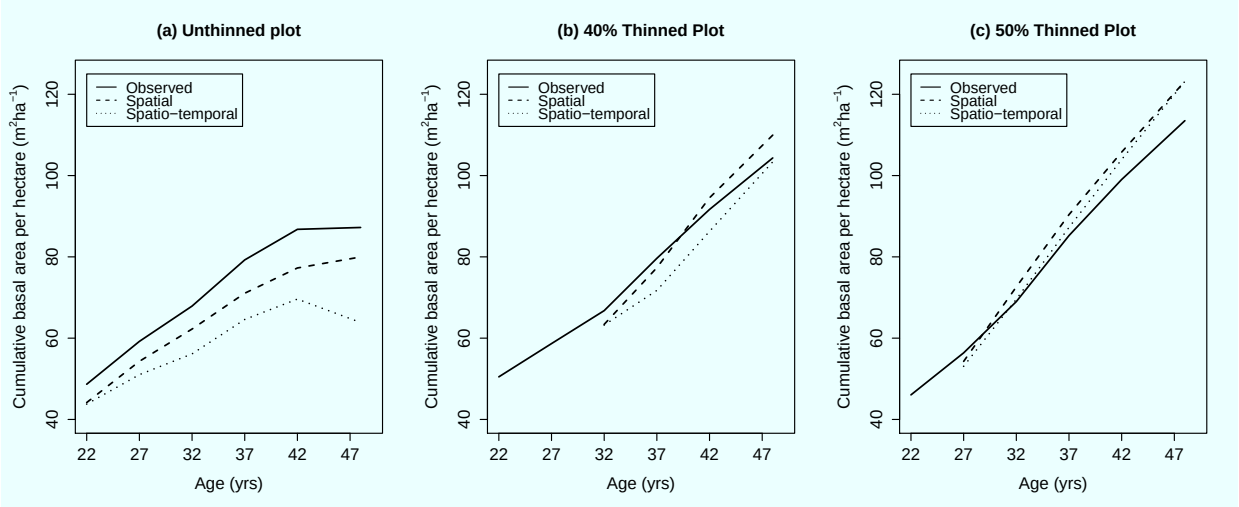


Figure 8: Observed, spatial predicted and spatio-temporal predicted cumulative basal area (CBA) per hectare (m^2ha^{-1}) after thinning for age from 22 to 47 years in (a) an unthinned plot, (b) a 40% thinned plot and (c) a 50% thinned plot.

679 List of Tables

Table 1: The root mean square prediction error (RMSE) for three models for an unthinned plot, for a 40% thinned plot and for a 50% thinned plot.

Model	Plot/Treatment	RMSE (cm)	$\hat{\lambda}$
1	Unthinned	5.774	0.384
2	Unthinned	4.206	0.384
3	Unthinned	4.205	0.384
		(2.426)	
1	40% Thinned	6.208	0.585
2	40% Thinned	4.266	0.585
3	40% Thinned	4.264	0.585
		(3.087)	
1	50% Thinned	6.208	0.747
2	50% Thinned	3.969	0.747
3	50% Thinned	4.088	0.747
		(2.924)	

Model 1. Box-Cox regression with repeated measures analysis.

Model 2. Spatial regression kriging (LOOCV).

Model 3. Spatio-temporal regression kriging (LOOCV using other trees and (LOOCV using previous measurements on a tree plus other trees)).

LOOCV: Leave-one-out cross-validation.

Table 2: Spatial variogram model parameters of the Box-Cox regression model residuals by treatment with all time points combined.

Plot/Treatment	Nugget	Sill	Range (m)
Unthinned	0.053	0.08	1.7
40% thinned	0.045	0.52	5.8
50% thinned	1.600	1.85	3.7

Table 3: Spatio-temporal variogram model parameters of the Box-Cox regression model residuals by treatment.

Plot/Treatment	Component	Nugget	Sill	Range	Anisotropy
Unthinned	Space	0.023	0.056	1.24 m	
Unthinned	Time	0.000	0.006	15 yr	
Unthinned	Space-Time	0.019	0.047	55 m	0.02 m/yr
40% Thinned	Space	0.000	0.196	1.33 m	
40% Thinned	Time	0.000	<0.001	10 yr	
40% Thinned	Space-Time	0.196	0.321	4.73 m	0.19 m/yr
50% Thinned	Space	0.591	0.001	1.86 m	
50% Thinned	Time	0.000	0.121	10.2 yr	
50% Thinned	Space-Time	0.482	1.145	3.1 m	0.18 m/yr

Table 4: Observed, spatial predicted, and spatio-temporal predicted cumulative basal area (CBA) (m^2ha^{-1}) per hectare after thinning for three plots at the last time point, in 1997.

	Initial thinning treatment		
Cumulative basal area (m^2ha^{-1}) per hectare after thinning	0% removal	40% removal	50% removal
Observed	87.23	104.33	113.50
Spatial predicted	79.98	109.95	123.20
Spatio-temporal predicted	71.08	103.38	123.17