Survival of the Fattest?
Self-thinning among Trees

Report 3, Department of Mathematics and Physics,
The Royal Veterinary and Agricultural University

Jens Lund*

April 29, 1998

Abstract

We consider data from an even-aged, unthinned Sitka spruce experiment in Denmark. The data is described in detail and the self-thinning process is modelled by a survival model for the discrete survival times. The survival model is based on Cox's proportional hazards model and allows for spatial dependence among the trees. The concept of competition indices is discussed at some length.

The results show that the small trees have a higher risk of dying than the large trees and that Hegyi's competition index based on basal area is a significant covariate in the model. The higher the competition index is, the higher is the risk of dying. Finally, there was also a significant dependence on the trees' positions in the experiment.

Preface

This report is part of a PhD course on forest biometrics under supervision of Jens Peter Skovsgaard, The Danish Forest and Landscape Research Institute (FSL). The course description is included in Appendix A.

1 Introduction

This report is an analysis of data from an even-aged, unthinned Sitka spruce experiment. The dataset includes breast height diameters at several time points as well as the positions of the individual trees. Several things are of interest, for example:

- The self-thinning process. As time goes by some of the trees die. A statistical description of this mechanism provides a significant reference for thinning strategies and for models for managed forest stands.
- The growth of the trees. The diameters were measured several times, so we have data on the development of the tree diameters over time. A model for this process would cast light on the development of even-aged forest stands and it would be of special interest to model the spatial competition between the trees.

*Address: The Royal Veterinary and Agricultural University, Department of Mathematics and Physics, Thorvaldsensvej 40, DK – 1871 Frederiksberg C, Denmark, e-mail: jlund@dina.kvl.dk.

- The spatial distribution of the diameters at a fixed time point.
- During the observation period for the experiment a minor gap was created in the southern part of the experiment due to windfall. It could be of some interest to quantify the spread of the windfall.

The time for this project did not allow me to study all these points, so I have studied only the self-thinning process. Due to data limitations this study should be considered a pilot project that may guide further, more comprehensive and detailed analyses.

In Section 3 we describe a discrete survival model with time dependent covariates. Among the covariates is competition indices that model spatial dependence among the individuals. The competition indices is based on the relative size of the trees. As it turns out, such competition indices might be the same as a description of the distribution of the tree sizes at a fixed time point.

This report is aimed towards both foresters and statisticians. As I am a statistician myself, some forest related comments are probably banal. On the other hand some of my points might be understandable by statisticians, but too technical for the average forester. I have tried to mark some of these points in the text.

In Section 2 we describe the data in detail and comment on graphs in Appendix B. The size of Section 2 reflects the amount of time I have spent on finding detailed descriptions of the data and getting familiar with the data. Section 3 is a description of the self-thinning model, and the results from the analyses are reported in Section 4. A discussion and some concluding remarks can be found in Section 5.

2 Description of Data

This section describes the data in various ways. Section 2.1 is an introduction to the experiment and Section 2.2 gives a detailed description of the dataset. The information in this section is mainly taken from [Sko97b] and the field note books with detailed information on each measurement. In order to give a sense of the dataset, a large number of graphs are shown in Appendix B. Section 2.3 contains comments on the graphs. The graphs are in the appendix, rather than in the main text, because they are best reproduced at a large scale and they would then clutter up the main text.

2.1 The Experiment

Experiment MBII is situated in the northern part of Jutland on Thy national forest district in "Nystrup dune plantation", compartment 437*e*. The experiment is conducted by The Danish Forest and Landscape Research Institute (FSL). This study comprises plot k, an unthinned plot of even-aged Sitka spruce. Plot k of the present experiment is located on exactly the same spot as the unthinned plot k of the previous experiment MB. Seeds for the present generation of trees originated from the previous experiment. The map in Figure 1 is reproduced from [Hen58, p. 296] and shows the old design of the southern part of experiment MB with plot k in the middle of the southern part of the experiment. The size of the plot is approximately 39m $\times$ 55m, which is 0.21ha. Around the plot is a border of a few meters with the same (non-)treatment as the experiment. As seen in Figure 1 the ground in experiment MBII slopes down towards the north-west corner, and the difference in elevation is 3 metres.

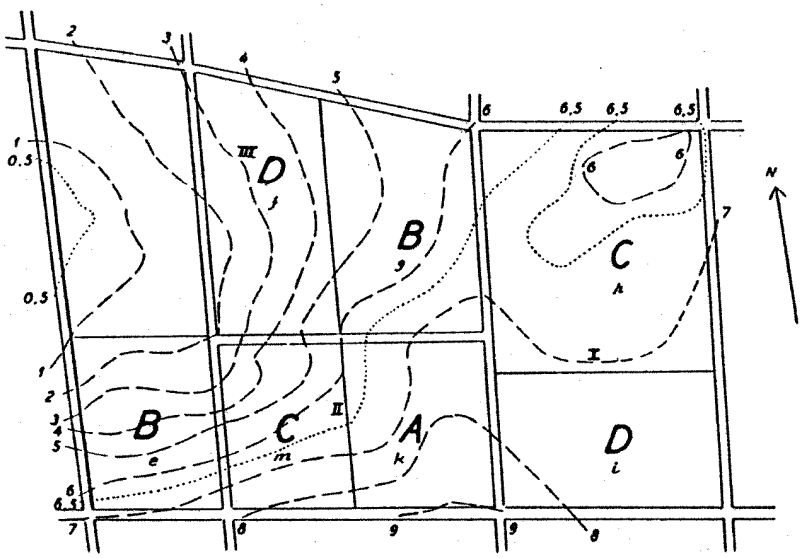

Figure 1: Map of experiment MB. Plot k of experiment MBII is placed at the old A-grade thinning in plot k.

The experiment was re-planted in the spring of 1957 with approximately half larch and half Sitka spruce, 927 trees in total. The larch trees were planted to prevent late frost damages in the Sitka spruce in the beginning. The larch trees were felled in 1972–1975, and the first measurements of the remaining 479 Sitka spruce trees were made in 1975. There is no further thinning in the experiment and the dead trees due to self-thinning are not removed from the experiment. The Sitka spruce stand is described as closed in the southern part and more open in the north-west corner in 1975. Although the terms "open" and "closed" here refer to the canopy, we get the same impression from Figure 13 on page 32 that shows a map with the diameters of the trees marked. At the first measurement in 1975 the basal area for Sitka spruce was 17.8m^2/ha and at the latest measurement in 1995 the basal area was 61.7m^2/ha. The corresponding number of living trees were 2228 trees/ha in 1975 and 1133 trees/ha in 1995.

2.2 The Measurements

The dataset consists of the positions of all the trees (including the larch trees), measurements of all the breast height diameters in 1975, 1979, 1984, 1990, and 1995, and measurements of some of the individual tree heights at the same occasions.

The rest of this section expands on this very short description, and supplies further details on the data collection.

The first 11 lines of the main data-file is displayed in Table 1 in order to make the following description of the dataset more comprehensible.

Each of the 927 trees are allocated a unique individual number that is used for reference purposes. The species of the trees are Sitka spruce (479 trees), larch (444 trees), and birch (4 trees). The birch and larch trees were felled and removed during 1972–1975.

For each tree we have the position in terms of row number and number within the row, as well as the coordinates in a usual xy-coordinate system. Figure 4 shows the positions and species in terms of row number and number within the row. We see that the main species Sitka spruce and larch are fairly homogeneously distributed. Figure 5

indiv	row	no	spec	dbh1975	dbh1979	dbh1982	dbh1984	dbh1990	dbh1995	x	y	remark1	remark2
1	1	1	SGR	34	0	0	0	0	0	NA	NA	""	""
2	1	2	SGR	81	87	0	94	0	0	-1.30	1.20	""	""
3	1	3	LAR	105	0	0	0	0	0	-1.42	2.57	""	""
4	1	4	SGR	140	164	0	187	215	243.5	-1.40	3.72	""	""
5	1	5	LAR	85	0	0	0	0	0	-1.52	4.95	""	""
6	1	6	SGR	137	158	0	171	176	0	-1.60	6.20	""	""
7	1	7	LAR	12	0	0	0	0	0	NA	NA	""	""
8	1	8	SGR	6	6	0	12	10	0	-1.52	7.66	""	""
9	1	9	SGR	104	132	0	178	230	268.5	-1.50	8.97	""	""
10	1	10	SGR	67	77	0	80	76	0	-1.50	9.40	""	""

Table 1: The first 11 lines of the data-file `mb2a.txt`.

shows all the positions in the xy-coordinate system. The x-axis runs approximately in the west-east direction and the y-axis runs in the south-north direction, and rows run roughly parallel to the y-axis. This means north will be in the up-ward direction on all the graphs. Compare also with the map in Figure 1. The unit on the axes is metres (m). The extent of the x position is from -40.1m to -1.3m and the extent of the y position is from 0.51m to 55.28m. As there are 24 rows, this means that the average spacing between original rows is $(-1.3\text{m} - (-40.1\text{m}))/(24 - 1) = 1.7\text{m}$. The original spacing within rows is on average $(55.28\text{m} - 0.51\text{m})/37.625 = 1.46\text{m}$, because the average number of trees in a row is $927/24 = 38.625$. This is in good agreement with [Sko97b], that states an estimated original spacing of 1.65m $\times$ 1.25m. Note that row number 1 has high x-values (close to 0m), whereas row number 24 has low x-values (close to -40m). The number within the row counts from low y-values to high y-values for the odd rows, and in the opposite direction for even rows.

The position in the xy-coordinate system is missing for 41 trees because the stumps (or the trees) have decayed or disappeared from the experiment. None of these 41 trees has any associated remarks. The distribution on species is Sitka spruce (20 trees), larch (20 trees), and birch (1 tree).

We would like to define a position for these 41 trees based on their row number and number within the row. We consider two different situations:

- One or more trees with missing positions are at the end of a row. In this situation the trees are placed in continuation of the row at a distance as between the two previous trees in the row.

  (These two trees have in all but one case their positions measured. In the single problematic case, tree number 732, the tree is placed at the distance as between the two previous trees with measured positions. This seems reasonable from a plot of the positions.)

- Otherwise, one or more trees are standing in a row with at least one tree on each side with the position measured. In this case the trees are placed uniformly in the space between the nearest two trees within the row with measured positions.

The defined positions are marked in Figure 5 with filled boxes.

I believe this algorithm is reasonable, and because the trees whose positions are missing tend to be small it will probably not matter that much what positions they have. Anyway, the error made by having a minor error on the positions of a few trees is probably not greater than if the trees are completely left out from the analyses.

The diameters at breast height (dbh), i.e. 1.3m above ground level, are measured on all the trees in the spring of the years 1975, 1979, 1984, 1990, and 1995. The unit of

the measurements is mm. Of course the larch and birch trees, that were felled and removed at the latest 1975, are only measured in 1975. If they were felled before 1975 their stump were measured. The exact dates the measurements are taken are 25 May 1975, 22 May 1979, 15 May 1984, 11 January 1990, and 10 May 1995.

On 24 November 1981 windfall occurred in the southern part of the experiment together with a few trees in the middle that were exposed too. The placement of the 43 fallen trees are marked in Figure 6. Dbh measurements were taken on the fallen trees on 30 April 1982, and the fallen trees were removed from the experiment afterwards. Note that the border of the experiment were not harmed by the windfall, a common phenomenon according to foresters. By now there are self-sown birch in the southern part of the experiment.

In the data file (Table 1), the dbh measurements implicitly define the time of death for the tree. We define the time of death as the time of the first measurement occasion the tree is not measured. Special care must be taken with the trees that are hit by the windfall in 1981 or still are alive at the last measurement in 1995. We only know that the trees survived longer than the windfall or 1995, respectively. In a survival analysis terminology they are "censored" at these events. We assume that the 43 trees in the windfall are censored at the measurement in 1984. This implies that the 43 trees were alive prior to the measurement in 1984, which might not be the case. However this seems the most reasonable to do. Table 2 shows the number of dead and censored Sitka spruce in the measurement years.

Year	'75	'79	'84	'90	'95	Total
Dead	1	2	9	66	115	193
Censored	0	0	43	0	243	286
Total	1	2	52	66	358	479

Table 2: Number of dead individuals, Sitka spruce.

In the same years as the dbh measurements were taken some of the heights were measured too. For a start 53 trees were measured in 1975, whereas only 30 trees were measured in each of the years 1979, 1984, 1990, and 1995. In total, height measurements involves 54 different trees, so the 30 trees were in general chosen among the original 53 trees. The rule is to use a tree as a height tree as long as it is alive. When a tree dies, another one of the original 53 trees are chosen to ensure that 30 height measurements were taken. The height trees are distributed regularly in space as seen from Figure 7. The unit of the height measurements is dm. We will not use the height measurements in the survival model in Section 3, but we comment on some graphs in Section 2.3.

Two remark fields in the data-set are used for additional information about the trees. The remarks are dead (9 trees), fork (41 trees), resin (1 tree), great spruce bark beetle (1), and bark peel by deer (1 tree). Note that only the first part of a forking tree is marked as such and the following parts are recognized by having the same position. Further note that two trees with the same position have the remark "not fork". The second remark field is only used four times, and in all cases to indicate a dead fork. The 13 trees with the remark "dead" are all measured at the last measurement in 1995, and the interpretation of the remark is that it will not be measured at the next measurement time. However this may also be true for a lot of other trees, so this information is not really useful and will thus not be used.

In the analyses below forks are not treated in any special way, but rather as two (or more) "independent" trees that happen to have the same position. This is quite reasonable. It is e.g. perfectly possible that one of the forks die before the others do. The two trees at the same position, that are not forks, are treated in the same way as if they were forks. This means they get no special attention. Of course, forks are not independent, but this approach means that the dependence (competition for light, water, etc.) among forks are modelled as the competition among all other trees.

2.3 Graphs

In this section we comment on some graphs of the raw data. We do not try to model the data in this section, but rather make unsophisticated observations from the graphs. The number of graphs is rather large so they are displayed in appendix B.

The graphs in the appendix are organized as follows:

- Figure 4–7 all regard the positions of the trees. We have commented on these graphs previously in Section 2.2.
- Figure 8 and 9 show the longitudinal development of the diameter measurements.
- Figure 10–12 display graphs of the diameter distribution.
- Figure 13–17 are maps of the experiment with circles that have a diameter proportional to the diameter of the tree placed at the position of each tree. Trees that have died before the measurement are marked by small filled circles. Note that the largest circle in each graph has the same size in all the graphs and that the scale of the circles are different from the scale on the axes.
- Figure 18–21 are maps similar to the previous ones, but with the diameter of the circle proportional to the diameter increment of the tree[1]. As before, note that the largest circle in each graph has the same size in all the graphs.
- Figure 22 is a plot of the increments versus the diameters.
- Figure 23 is a collection of plots of height versus dbh and log(dbh).

The plot of the longitudinal development in Figure 8 can be very disorderly to look at, so in Figure 9 all the lines are displayed exactly once in one of ten displays. The lines in the plot stop at the last measurement, so when the trees are measured the last time, it looks almost like a vertical line of line ends. When looking at the trees that die two things can be noted. First, the trees that die tend to be small, and second, they seem to have very small increments. It can of course be discussed whether the trees have small increments because they are dying or they die because they are small and have small increments.

The larch trees in the experiment were present to help start the Sitka spruce culture. From Figure 10 it is immediately clear that the larch trees are in fact larger than Sitka

[1] This does not imply that the area of the circle is proportional to the increment in the basal area. One might argue that the eye is more focused on changes in the area of the circles than in changes in the diameter, and further that the area of the circle should be proportional to the diameter increment. One might also argue that the basal area increment is more biological relevant and should be used instead. The variations are endless, and the graphs show more or less the same picture irrespective of variable chosen.

spruce in 1975. This means that they, at least in the beginning, grow faster than Sitka spruce.

Figure 11 and 12 compare the size distributions for the Sitka spruce trees in the measurement years. The boxplots[2] in Figure 11 are a simpler way to illustrate the size distribution than the histograms in Figure 12. It is seen that the size distribution is right skewed in all the years. This is expected when the trees are small (after all, they all start at 0 at the same time), but it is also the case when the trees become larger. This is probably a kind of starting effect: the trees that start well are most likely to fare well for the rest of their lives. The reason that a particular tree is better off from the beginning might be pure chance or a difference in the genes of the trees. The histograms in Figure 12 show that the distributions are in fact very broad, particularly for the later years. The size distribution might be described as approximately uniform on an interval with a heavy right tail attached. This is also the impression from the boxplots in Figure 11 where it is seen that the first quartile has approximately the same size as the distance from the first quartile to the median and as the distance from the median to the third quartile, but the distance from the third quartile to the right end of the data is larger than this common distance. In [Sko97a], J. P. Skovsgaard outlines in Section 6.3.2 a theoretical development of the size distribution and he uses this as a part of his hypothesis 2 on page 52. The theoretical development of the size distribution is left skewed $\rightarrow$ symmetric $\rightarrow$ right skewed $\rightarrow$ symmetric. The present dataset cannot confirm this development. This might be because we do have diameter measurements for a 20 year period only and thus do not cover the entire life span of the trees. J. P. Skovsgaard [Sko97a, p. 187] also rejects the hypothesis for stands like the present, although he in contrast talks about *left* skewed distributions during the whole life span.

One striking feature when comparing the histograms in Figure 12 is that all the distributions, except that from 1995, have a number of small trees, i.e. trees with diameter below say 10cm. All these small trees were declared dead in 1995. It could of course very well be that they died. But as seen from Figure 8 it looks like these trees have had about the same size for a long period, and it is amazing all of them die in the same time interval.

The maps in Figure 13–17 tell something about the spatial distribution of the sizes. As already noted the map from the start of the experiment in 1975 (Figure 13) shows that the northern part of the experiment is more open than the southern part. This impression remains valid during the whole period. Note that from 1984 and on the main part of "dead" trees in the southern part of the experiment is the trees in the windfall. These were removed. If we look at the positions of dead trees in 1990 and 1995, Figure 16 and 17, it seems like the dead trees are more likely to be in the east side of the experiment than the west side. This is confirmed by the results in Section 4 from the survival model in Section 3. It is seen from the plots that the small trees are more likely to die than the large trees.

The maps in Figure 18–21 of the increments suggest that in the beginning, i.e. 1979 and 1984, the increments are larger in the more open northern part than in the southern

[2] A boxplot tells some basic facts about the distribution in a way that makes it easy to compare several distributions. The horizontal line in the interior is located at the median of the distribution. The box starts at the first quartile (25%) and stops at the third quartile (75%). The whisker extending from the top of the box goes to the first data value below the median+1.96×"the interquartile distance", where 1.96 is the 0.975 quartile of a $N(0,1)$ distribution. The whisker at the bottom is defined in a similar way. For data having a Gaussian distribution, approximately 99.2% of the data falls inside the whiskers. Data points which fall outside the whiskers are indicated by horizontal lines.

part. Later on, 1990 and 1995, the increments seem to be more uniformly distributed. Figure 21 of the increments up to 1995 indicates that a few trees have very large increments compared to the other trees because there are many small circles present. In fact, Figure 9 of the longitudinal development shows that the increments in the period up to 1995 are smaller than in the previous periods and that a few trees despite this have large increments in the last period. One of the main conclusions in [Sko97a, p. 202] is that unthinned Sitka spruce stands is good at differentiating the diameter sizes. A comparison between the maps of the sizes of the trees with the maps of the increments show that the large trees also have the largest increments. This is probably more clear from Figure 22 that directly shows the increments and diameters plotted against each other. When comparing the plot of the dead trees marked in 1990 with the plot of the increments in 1984 it is seen (once again) that the trees that die are the trees with small increments. It is quite clear from the plot of the increments that more trees have negative or zero increments in the eastern part than the western part of the experiment. This can also be seen from the plots in 1995 (dead trees) and 1990 (increments), although not as noticeable.

Henriksen [Hen81, p. 14, 56] argues that the diameter-height-regression height $= \alpha + \beta \log(\text{dbh})$ is a simple but very useful model. Figure 23 is a set of plots of height versus dbh and log(dbh). At first sight it looks like the simple regression height $= \alpha + \beta \text{dbh}$ would be a good fit to the data in 1975 and 1979, whereas the regression height $= \alpha + \beta \log(\text{dbh})$ would be best in 1990 and 1995. In 1984 either regression would fit the data. This means that there is a development in time: the relation between dbh and height changes over time. Henriksen [Hen81, p. 21] mentions several aspects of the development in time, but not that the regression with log(dbh) should be inadequate. I haven't found any further comments in [Hen81] on this, but have been told that the regression with log(dbh) is often not adequate in unthinned stands.

3 Self-thinning — a Discrete Survival Model

3.1 Introduction

In this section we make a model for the self-thinning. This is in fact a model for survival of the trees. The survival times for the trees are continuous, but the survival times are only observed to be in some discrete intervals between the inspections and we thus have interval censoring. The stand is even aged, so all the trees are censored at the same ages and we have no hope of estimating the continuous life time distribution. In the following we make a discrete survival model with spatial dependence and we end up by modelling the discrete survival times with a grouped version of Cox's proportional hazards model.

Section 2.2 describes among other things how the discrete dead and censoring times in Table 2 are inferred from the measurements of the diameters. We will use all the living Sitka spruce trees in 1975 as our population and Table 2 shows that this is 478 trees.

Sections 3.2–3.3 are an overview of various models for discrete survival times. We start by introducing some notation (Section 3.2) and go on to describe models where the discrete time hazard is modelled (Section 3.3). Section 3.4 and Section 3.5 discuss in detail how time dependent covariates should be used and how we will allow for spatial dependence through competition indices.

3.2 Notation for Survival of One Tree

Let $Y \in \{1, \ldots, k\}$ be a discrete survival time with distribution $P(Y = j) = \pi_j$ for $j \in \{1, \ldots, k\}$. Define the discrete distribution function $\gamma_j = P(Y \leq j) = \pi_1 + \cdots + \pi_j$ and the discrete hazard $\lambda_j = P(Y = j | Y \geq j) = \frac{\pi_j}{\pi_j + \cdots + \pi_k} = \frac{\gamma_j - \gamma_{j-1}}{1 - \gamma_{j-1}}$. Note that $\pi_j = \lambda_j \prod_{i=1}^{j-1}(1 - \lambda_i)$ and $P(Y > j) = 1 - \gamma_j = \prod_{i=1}^{j}(1 - \lambda_i)$.

If we have a vector of covariates $x' = (x_1, \ldots, x_d) \in \mathbb{R}^d$ in addition to the survival time Y we will write $\pi_j(x)$, $\gamma_j(x)$, and $\lambda_j(x)$ for the point probabilities, distribution function and the discrete time hazard given the covariate x. This means e.g. that we have the relation

$$\pi_j(x) = \lambda_j(x) \prod_{i=1}^{j-1}(1 - \lambda_i(x)). \tag{1}$$

3.3 Modelling of Discrete Time Hazards

The discrete time hazard $\lambda_j(x)$ is often modelled when we consider discrete survival times. Sheike and Jensen [SJ95, Sec. 2] outlines several such approaches. One advantage of modelling $\lambda_j(x)$ is the easy interpretation in a survival analysis context and that we can use standard software for generalized linear models (GLM) for estimation. Both of these aspects use the connection (1) between π_j and λ_j. In the survival analysis context here it is natural to interpret this relation as a product built stepwise as time goes by since one can only reach a state (time point) by going through all the former states (time points). These models are also called "sequential models" in [FT94, Sec. 3.3.4]. Estimation is easy with standard software for GLM because the likelihood (1) for one individual can be interpreted as the likelihood for a "fake" dataset of binomial variables $z_1 = 0, \ldots, z_{j-1} = 0, z_j = 1$, so that $\pi_j = \prod_{i=1}^{j} \lambda_i^{z_i}(1 - \lambda_i)^{1 - z_i}$ as outlined in [FT94, p. 322]. Censoring can of course also be handled by letting all the y variables in the fake dataset take the value 0.

We will now look at some specific choices of the link function for $\lambda_j(x)$. The link function says how $\lambda_j(x)$ depends on the covariate vector x. The sign of the parameter $\beta \in \mathbb{R}^d$ in the following is chosen so that larger values of the components of x give higher mass to large values of Y provided $\beta > 0$.

Proportional hazards Assume that there is an underlying continuous survival time with continuous time hazard $\lambda(t|x) = \lambda_0(t)e^{-x'\beta}$, which is Cox's proportional hazards model. Let dt denote the length of a small time interval. Then the interpretation of the hazard is that $\lambda(t|x)\, dt$ is the probability of dying before time $t + dt$ given the individual has survived to time t. The survival function is $S(t|x) = \exp(-\Lambda_0(t)e^{-x'\beta})$ with $\Lambda_0(t) = \int_0^t \lambda_0(s)\, ds$. When we make a grouped version $Y \in \{1, \ldots, k\}$ of a variable with this distribution according to intervals $[0, \theta_1],]\theta_1, \theta_2], \ldots,]\theta_{k-1}, \infty]$, we get $\gamma_j(x) = 1 - S(\theta_j|x) = 1 - \exp(-\Lambda_0(\theta_j)e^{-x'\beta})$ and $\lambda_j(x) = 1 - \exp(-e^{-x'\beta}(\Lambda_0(\theta_j) - \Lambda_0(\theta_{j-1})))$. This means that the complementary log-log transform[3] $\operatorname{cloglog}(\lambda_j(x)) = \log(-\log(1 - \lambda_j(x))) = \hat{\theta}_j - x'\beta$ of the hazard $\lambda_j(x)$ is linear with $\hat{\theta}_j = \log(\Lambda_0(\theta_j) - \Lambda_0(\theta_{j-1}))$. This is e.g. described in [FT94, p. 318].

[3] The complementary log-log transform is the inverse of the distribution function $F(t) = 1 - \exp(-e^t)$ for the extreme-minimal-value distribution.

Logistic model The logistic model uses the logit link[4]

$$\text{logit}(\lambda_j(x)) = \log \frac{\lambda_j(x)}{1-\lambda_j(x)} = \tilde{\theta_j} - x'\beta.$$

See references in [FT94, Sec. 9.2] and [SJ95].

Log-link $\log(\lambda_j) = \tilde{\theta_j} - x'\beta$. See reference in [SJ95]. This is a proportional hazards model for the discrete time survival time distribution. A problem is that we must have unnatural restrictions on the parameter space, because $\tilde{\theta_j} - x'\beta$ must be < 0 in order for λ_j to be in $]0, 1[$.

We use the grouped version of Cox's proportional hazards model in the following. The choice of this model is primarily for its appealing and easy interpretation.

3.4 Time Dependent Covariates

We go into detail about the choice of covariates in this section as we want to have time dependent covariates and spatial dependence among the individuals. The use for time dependent covariates is obvious because we suspect the probability of dying to depend on the current size of the tree, which develops through time. Furthermore we suspect the trees to have a degree of inter-dependence. If two large trees are standing very close to a small tree we would expect the small tree to have a higher risk of dying than if the two large trees were not present. So we want to model some kind of spatial dependence between the trees.

The discrete time hazard models in Section 3.3 are very easily modified to allow time dependent covariates. Assume that for each time point $1, \ldots, j$ up to and including the time of death j we have a covariate vector, $x_1, \ldots, x_j$. We now exploit the Markov structure in equation (1) by using the likelihood function $\pi_j(x) = \lambda_j(x_j) \prod_{i=1}^{j-1}(1 - \lambda_i(x_i))$, where in each step we condition on the present value of the covariate, but still use the same link function as before for the hazard λ_j. In our application on trees it seems natural to use the dbh measurements at the *previous* measurement as covariate in the regression for survival in the present period because the size of the tree affects what comes in the following period. A sketch of this setup is in Figure 2.

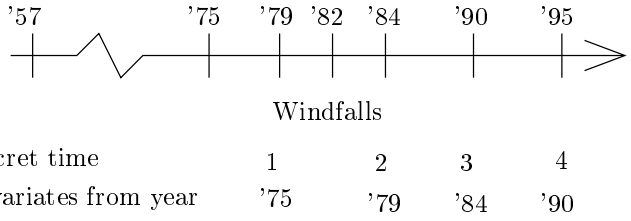

Figure 2: Sketch of discrete times and covariates used.

We have only considered models for one individual so far, and we will now go on and describe a model for all the trees in the stand and allow for dependence between the trees. In the classical survival analysis context dependence between the individuals have e.g. been modelled by frailty models[5]. (A recent reference is [Pet96].) However,

[4]The logit transform is the inverse of the distribution function $F(t) = \frac{e^t}{1+e^t}$ for the logistic distribution.

[5]In short, a frailty model is a model with an unobserved latent variable.

this approach does not seem natural here as the dependence among individuals depend on the distance between them and not on any natural groups of individuals. We allow for dependence through the covariate process instead.

We will build the model stepwise through discrete time. Let S denote the set of trees and let $t = 1, \ldots, k$ be the discrete time points considered. The special time point $t = 0$ denotes an imaginary time point before the first inspection. Introduce the set of living trees S_t at time t. We assume that all trees are alive before time $t = 1$ so $S_0 = S$. Further note that $S_t \supseteq S_{t+1}$. For $s \in S$ and $t \in \{0, \ldots, k\}$ we put $N(s,t) = N_s(t) = 0$ if tree s is alive at time t and $N(s,t) = N_s(t) = 1$ if tree s is dead at time t. This means $S_t = \{s \in S \mid N_s(t) = 0\}$. Define the covariate process X as $X(s,t) = X_s(t)$, the measurements on tree $s \in S$ at time t. By $X(t)$ we denote the collection of all measurements at time t, $X(t) = (X(s,t))_{s \in S}$, and if a tree is dead at time t the dead status is the measurement.

We now assume that conditionally on the covariates $(X(s,t-1))_{s \in S_{t-1}}$ and the dead/alive status at the previous measurement the survival of trees in the following time period are independent events. The important point here is that we allow the survival probability of tree s to depend on the measurements and the status of the other trees $S \setminus s$. The likelihood is now

$$L = \prod_{t=1}^{k} \Big(\prod_{s \in S_{t-1} \setminus S_t} \lambda_{s,t}(X(t-1)) \prod_{s \in S_t} \big(1 - \lambda_{s,t}(X(t-1))\big) \Big). \tag{2}$$

Here $\lambda_{s,t}(X(t-1))$ denotes the hazard of tree s at time t with covariates $X(t-1)$. With this notation the covariate x_i at time i is $X(i-1)$.

For each tree s we will use the measurements on the tree itself and some measure of the competition from the neighbouring trees as covariates. In Section 4 we mention the covariates based on the single tree and in the next Section 3.5 we discuss the measures of competition from the neighbouring trees.

It should be noted that the present setup without problems can be generalized to a continuous time setup. (This paragraph might be a bit technical.) The framework of counting processes as described in [ABGK93] is natural to use in the generalization. The counting process setup does not handle simultaneous deaths easily, but this problem could be solved if necessary. The choice of covariates measured at the previous inspection is quite obvious in the discrete time setting here. However, in a continuous time setting the analogue is that the covariate processes should be predictable in the relevant filtration. It is quite easy to check, that if we in the proportional hazards model in Section 3.3 have a predictable time dependent covariate that is piecewise constant between the grouping time points, then we get a model that is linear in the parameters with a cloglog transformation of λ_j. This model is exactly of the above type. This means that we can regard the discrete time hazard model with cloglog-link as a grouped version of a Cox regression model with time dependent covariates.

3.5 Definition of Competition Index

We would like to include a measure of the size of a tree compared to the other trees surrounding it as a covariate in the regression model. This could tell something about the competition between the trees. V. K. Johansen [Joh96] reviews several such measures also known as *competition indices* (CI) in order to make a growth model spanning

over several periods. Tomé and Burkhart [TB89] also reviews a rather large number of CIs and suggest some modified indices and compare how good the indices predict growth. Pukkala [Puk89] compares two different approaches to the definition of a CI and takes into account the direction of the competition. In this paper we focus on a competition index of the type suggested by Hegyi [Heg74], which often is rated as one of the best, rather than compare several competition indices. We end this section with some more general and technical remarks about the definition of competition indices.

Let $A_s(t)$ denote the basal area for tree $s \in S$ at time t. We will leave out the dependence on t in the following since all variables will regard the same time point. At each time step we recalculate the competition index based on the measurements at the previous time point. The competition index suggested by [Heg74] is

$$\mathrm{CI}_s = \sum_{\tilde{s} \neq s \,:\, \mathrm{d}(s,\tilde{s}) \leq 4\mathrm{m}} \frac{A_{\tilde{s}}}{A_s \, \mathrm{d}(s,\tilde{s})}$$

where $\mathrm{d}(s,\tilde{s})$ is the distance between tree s and tree $\tilde{s}$.

We modify this CI slightly. First of all we have trees (e.g. forks) at a distance from each other of 0m that would cause the CI to take the value ∞. Some trees are also very close so they cause an unrealistic high CI. These problems are alleviated by defining the minimum distance in the CI to be 0.5m. Second, some small trees also get an unrealistic high CI. In order to avoid that a minority of the trees (approximately 15%) make the distribution of CI highly skewed we impose the restriction that the factor $A_{\tilde{s}}/A_s$ is at most 16. With my modifications the CI is

$$\mathrm{CI}_s = \sum_{\tilde{s} \neq s \,:\, \mathrm{d}(s,\tilde{s}) \leq 4\mathrm{m}} \frac{\min(\frac{A_{\tilde{s}}}{A_s}, 16)}{\max(\mathrm{d}(s,\tilde{s}), 0.5\mathrm{m})}$$

with the sum ranging over living trees only.

Johansen defines in [Joh96] the competition index as the sum over the n living neighbour trees nearest to tree s. I find my approach that is also used in [Heg74] more natural, but it will probably not make any great difference in this experiment where the trees are placed quite regular.

In the same way as above, we define a competition index based on the diameters instead of the basal area. In this case we take the maximum fraction between the diameters to be 4. The original definition of the CI in [Heg74] was based on the diameters.

Figure 3 is two plots of the CI based on basal area in 1975. As seen from the first plot the distribution of CI is heavy tailed. The second plot shows the CI versus the diameters and it confirms that a large CI is most predominant among small trees. It may be noted that from several plots like Figure 3 it seems like the distribution of the CI is approximately the same over time, and that the distribution of the CI based on the diameters have a quite similar shape, although a different scale and not quite as heavy a tail.

We will now go on to some more general and technical considerations about competition indices. This discussion was initiated by Antti Penttinen in some remarks to a talk given by me at "workshop: spatial statistics and GIS" in Gothenburg, November 25–26, 1997. The discussion continued in private communication [Pen97].

Stoyan and Grabarnik [SG91] defines *energy marks* for a Gibbs point process and show certain moment properties of these energy marks. The results are also mentioned in [SKM95, p. 180, 191], which is an overview of stochastic geometry. In our

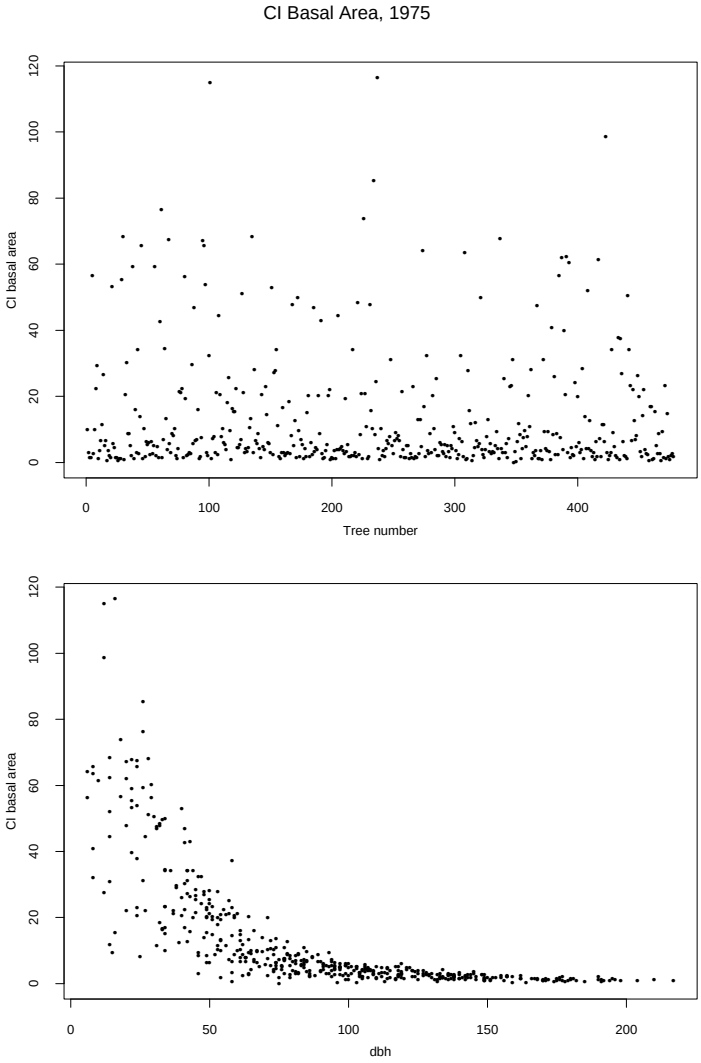

Figure 3: Plots of CI based on basal area in 1975. The upper plot shows CI versus the number of the tree, so the vertical distribution of points give an impression of the distribution of CI. The lower plot shows CI plotted against dbh.

situation the positions are fixed in advance and we will condition on the positions, so we will instead of point processes need the theory for Markov random fields (MRF). The definition of the energy marks can easily be generalized to MRF and marked Gibbs processes, and essentially the energy mark is a bijective transformation of the conditional probability density at each node (tree) given all the other nodes (trees). This figure is of course a measure of the "stress" on each node caused by the neighbours or in other words the relative size of a tree compared to the neighbours. Because we compare with the neighbours only, the relative size is in a local sense and not in a global sense as e.g. the quartile corresponding to the size compared to the size of all the trees. Intuitively this conditional probability is an "optimal" choice of CI and an interesting problem is to make precise the meaning of the word "optimal". In our case we look for a linear predictor in a specific GLM-model.

The drawback of this approach is that a reasonable distribution for the MRF consisting of the tree diameters at a specific time point must be specified and estimated. In this context it is natural to suggest a Gibbs type distribution specified by a set of neighbourhood relations and a pair potential function θ. The neighbourhood relations can e.g. be chosen as the natural "lattice" put on top of the set of trees[6]. The likelihood for such a distribution is

$$p(X) = \frac{1}{Z} \exp\left(\sum_{s \sim \tilde{s}} \theta\left(X_s, X_{\tilde{s}}, \mathrm{d}(s, \tilde{s})\right) \right)$$

where Z is the normalizing constant, the sum is taken over all pairs of neighbours, and the pair potential θ is allowed to depend on the distance between the two nodes besides the values at the two nodes of course. The dependence on the distance compensate for the imperfectness of the "lattice". The pair potential function might be estimated by non parametric methods or have a parametric form. Non parametric estimation of pair potential functions is e.g. considered in paper 2 in [Hei97], namely [HP95].

A point of initial confusion for my self was the dependence on distance but not e.g. direction. It is easy to verify that it is not reasonable that the covariance function between two nodes (trees) does depend on the distance between the two nodes only: two trees at distance, say, 3m cannot have the same covariance[7] irrespective of whether there is a tree in between. However, it seems more natural that the pair potential function does depend on the distance only. The pair potential function measures the "interaction" between two neighbours at short distances whereas correlation might very well exist at larger distances. Even though the pair potential does only depend on distance the covariance function also depend on the positions of the points because of the neighbour relations.

The above approach suggest that we must make a model for the simultaneous distribution of the diameters at a fixed time point. On the other hand we defined a competition index, that was a sum over neighbours with each term depending only on the actual tree, the neighbour, and the distance between them. In this way we have actually defined a MRF for the diameters. The energy for a specific node (tree) given the other nodes is the CI!

[6] The lattice is not a perfect lattice, e.g. because the number of trees in a row vary. We use a neighbourhood structure that looks as much as possible as a lattice.

[7] Wälder and Stoyan [WS96] outlines several problems when variograms are used to make models in "point process statistics" instead of the usual context of geostatistics. These problems are the same in the MRF setting. The problems arise especially when there is competition between the points.

These considerations seem to imply that the problem of defining a good CI is the same as making a good model for the simultaneous distribution of the diameters at a fixed time point and then find the conditional probabilities. This idea has not been pursued further here.

4 Results

In this section we present the results from the analyses of the model with the likelihood function (2) and a cloglog-link for the discrete time hazards $\lambda_{s,t}(X(t-1))$ as described in Section 3.

We use the diameter (dbh), the basal area (ba), and the competition indices based on the diameter (cidbh) and on the basal area (ciba) as covariates in our model. In Section 2.3 it was noted that Figure 16 and 17 suggest that the dead trees are a bit more likely to be in the east side of the experiment. For this reason we also include the coordinates x and y as well as their product[8] as covariates.

The results from the analysis show the consistent picture that the variables dbh, competition index based on basal area, and x and y coordinates are significant, whereas the variables basal area and competition index based on dbh are non significant. Likewise the product xy of the x and y coordinates can be removed from the model.

The test probability of removing the product xy from the full model is 53%. It seems natural that one of dbh and basal area, as well as one of competition index based on dbh and basal area should be in the final model. The test probabilities in Table 3 support the impression that the model with dbh and competition index based on basal area gives the best description of the data.

	ciba	cidbh
basal area	12%	$<$0.1%
dbh	50%	10%

Table 3: Test probabilities for the model including the variables in the table as well as the x and y coordinates against the full model.

Table 4 shows the parameter estimates. The time parameters are not that interesting per se, but it is worth noting that the values get larger as time goes by. This is because the probability of dying increases with time which is immediately seen from Table 2. The interpretation of the parameters for the covariates is best done in the underlying Cox proportional hazards model.

Table 5 shows for the four covariates the factor that should be multiplied on the continuous time baseline hazard with a certain increase in the variable and the confidence interval for these values. These values are based on Table 4, and e.g. the value for dbh is $\exp(-50*0.023) = 0.32$ and the confidence interval is the confidence interval for the parameter transformed with the exponential function. It is a part of the underlying Cox model that this factor depends only on the increase and not on the current value of the covariate.

It is seen that whenever the diameter is increased by 5cm the hazard is only one third of what it was previously. This means that the larger the tree is the smaller is the risk

[8]The product xy of the coordinates can be thought of as a "pseudo covariate" that models a sort of dependence between the x and y coordinates.

Variable	Estimate	Std. Error
time1	-3.90	0.80
time2	-2.29	0.57
time3	0.41	0.45
time4	2.94	0.49
dbh	0.023	0.0028
ciba	-0.024	0.0060
x	-0.026	0.0076
y	0.016	0.0064

Table 4: Parameter estimates.

Variable	Increase	Factor	Confidence interval
dbh	5cm	0.32	[0.24; 0.42]
ciba	10	1.28	[1.13; 1.43]
x	10m	1.30	[1.12; 1.51]
y	10m	0.85	[0.75; 0.97]

Table 5: Factor to multiply on the baseline hazard in the Cox proportional hazards model.

of dying. The hazard increases with almost one third when the CI is increased by a value of 10. In Section 2.3 it was noted that the dead trees seem more dominant in the east part of the experiment. This is confirmed by the estimate of factor for the x parameter — when we go 10m more to the east, then the hazard is increased by one third. The estimate of a decrease of the hazard by 15% when we go 10m north is not that obvious from the plots in Appendix B. On the other hand the test probability of removing the y variable is around 1.2% so this effect is not highly significant.

The choice of radius 4m in the definition of the competition index is somewhat arbitrary, although the discussion in the end of Section 3.5 suggests a relatively small radius. However, the same analyses as above were also carried out with a radius of 15m in the definition of the CI. The test probabilities and the parameter estimates are almost the same and the same variables are significant. However, there is a tendency that the estimates of the parameters for the two competition indices are about 1/3 of the value when the competition indices are defined with a 4m radius. This means that the factor above for the CI should be raised to a power of 1/3 and thus the factor will be closer to one.

5 Discussion and Conclusion

In the above analyses we have not talked about model checking and goodness of fit tests. The underlying Cox model make two assumptions:

- The hazards are proportional for all individuals.
- The hazard is log-linear in the covariates.

We cannot really check the first assumption in this data set since we do not have deaths on many time points. Table 2 shows that the vast majority of deaths fall in the

last two time intervals so in fact we almost only observe the dead time to be in one of three intervals.

The log-linearity would usually be checked by some plots, but it is not possible to make any good diagnostic plots because we only have one 0-1 observation for each combination of the covariates. Instead the model can be checked by introducing a factor with groups of the continuous covariates which we subsequently try to remove in a test. Thus the covariates dbh, ciba, cidbh, x, and y, are grouped into factors at 4 levels. The levels are chosen to divide the covariates at the quartiles[9]. The goodness of fit test that removes all the factors have a test statistic with value 34.2 with 15 df. This gives a test probability of 0.3% which would normally be regarded as significant and we must reject our initial model. However, it is not extremely significant and in a stepwise test with removal of the factors one after one, no one of the tests have a test probability below 2.6% if the tests are done in a "clever order". If the tests are done in a "random" order the lowest test probabilities are around 1% and there are several high ($>$10%) test probabilities. As already mentioned the distribution of the competition indices is highly skewed with a heavy right tail and this might also affect the behaviour of the regression estimates. All in all I believe that the results from the analyses are reliable, although further investigation of the behaviour of the model should be carried out. These investigations could be on the influence of the highly skewed distributions of the competition indices, and over-dispersion models as e.g. random effects models considered in [SJ95]. Unfortunately the time for this project do not allow me to pursue these important matters further.

We also assumed that the trees were independent in the following period given the current status. This assumption is obviously not fulfilled since there is an ongoing competition between the trees, but it suffices as a simple approximation. The validity of this approximation might perhaps be investigated by some sort of permutation test. In general the "level" at which some sort of independence is assumed might be investigated further together with the consequences for the model as a whole. In Rathbun and Cressie [RC94] they find in a growth model that it is satisfactory to describe the increments as independent, whereas Penttinen et al. [PSH92] find in two of three examples that the size of trees at a fixed time point can be described as independent. One of the two examples with independent sizes in [PSH92] is in a thinned plot.

In this report we use a model for the discrete time hazard to describe the discrete survival times. These discrete survival times could be considered as ordinal data and a common regression model for ordinal data is the McCullagh model as described in [McC80], [FT94, Sec. 3.3.1] or [MN89, Sec. 5.2.2]. These models are not suitable for our purpose because they do not allow the incorporation of time dependent variables in any obvious way, however.

It would have been interesting to study the effect of the dead trees on the living trees and in e.g. [TB89] the dead trees are included in the calculation of a special CI. But as noted almost all the trees died in the last two periods so it would just make sense in the last period and we have not gone further into this.

As seen in Figure 8 it seems like the diameter increment is approximately zero for the trees that die, so it is tempting to use the diameter increment as a covariate. This is not practical, however. We cannot use the increment in the time period going back to

[9]This is the reason the the covariate basal area is not grouped. Such a factor would be identical to the factor based on dbh since basal area is a monotone transformation of dbh.

the previous measurement because we must use the measurement at the present time point to find this number. This means that we use a measurement taken at the present time point to predict the course of the tree in the previous period. In mathematical terms the covariate process is no longer predictable. In the present data set we have relatively few time points and it is not realistic to use e.g. the diameter increment during the period between the former two measurements as a covariate.

We define the neighbours in the definition of the CI to be all the trees within a certain radius and as already noted this is probably not very different from taking the n nearest neighbours because the trees are placed regularly. The considerations in Section 3.5 suggest that this radius should be relatively small, but still the choice of 4m is somewhat arbitrary. It would be desirable to estimate this radius from data. Another idea would be to define several competition indices based on trees in different distances intervals, e.g. 0–2m, 2–4m, etc. Hegyi [Heg74] also uses a small distance of 10ft, whereas Rathbun and Cressie [RC94] uses a rather large distance of 30m or even larger in their definition of a CI.

In the analyses we have not worried about border effects. In the border of the plot, Sitka spruce of the same age and at the same spacing is planted, and we should somehow correct the competition index near the borders to reflect this. Note that near the borders the behaviour of the CI would depend on whether we included trees within a certain radius or we include the n closest trees. I don't think the question about border corrections is essential for the validity of the results.

The dependence of the hazard on the x and y coordinates is probably due to the slope of the ground. It might be that more of the "weak" trees in the low north and west part of the experiment died before 1975, so that more weak trees are present in the east and south part. The death of trees in the north and west part of the experiment can e.g. be caused by more frost in the lower part or differences in the ground water level.

The general considerations about competition indices in Section 3.5 suggest that a good competition index is approximately the same as specifying a good model for the sizes of the trees at a fixed time point. I have not found any references to this way of thinking of competition indices and it would be worth investigating further. With examples from forestry Penttinen et al. [PSH92] demonstrates several summary statistics and graphs in the context of marked point processes that say something about the spatial distribution of the sizes of trees. See also [WS96] for comments on the use of usual spatial models in a forestry situation.

In most of the models in [RC94] they treat the trees in three different size classes based on the diameter. We have not done this here and I suspect that they must do it this way instead of just looking at the diameters because they are not looking at even aged stands as I do.

The basic measurement on the trees is the diameter at breast height, dbh. We also use this measurement as basal area ($\propto \text{dbh}^2$) and could use it as volume raised to the power α, dbh^α. In a similar way the competition indices could be based on dbh^β for a general β. We found that we should use the dbh measurements and the CI based on the basal area. In [TB89] they make a growth model and conclude that the competition indices based on diameters always were superior to those based on basal area. The situation in Rathbun and Cressie [RC94] is a little different. They investigate an uneven aged stand and model both germinations, growth and deaths of trees. In the study of deaths they conclude that the competition from the neighbours does depend on their number

and distances only, and not the size of those neighbours. In the growth model they also conclude that a competition index defined on the basis of diameters and the distance in the same way as mine is the best. So these two references find that competition indices based on diameters are better than competition indices based on basal area, which I find the best. I cannot give any explanations for this, but only guess that it might be due to the different nature of the models; growth models versus survival model.

It is of interest to make general non linear regression models where we estimate the parameters α and β. This would be relatively straight forward for the α parameter, but computationally heavy for the β parameter since we need to recalculate the CI for each new value of β in an iterative estimation procedure.

In the definition of the CI we divide by the distance, but could as well divide by $\mathrm{d}(s,\tilde{s})^\gamma$. The case $\gamma = 2$ appears often in physics and Tomé and Burkhart [TB89] tries this among other variants, but they do not find conclusive evidence on which value of γ to use so they stay with $\gamma = 1$. In [RC94] they also find the use of $\gamma = 1$ to be satisfactory. We could estimate the γ parameter in the same way as the β parameter.

An alternative to the non linearity in α is to use log(dbh) as a covariate which would cause α to enter the model in a linear way.

The growth model by Pukkala [Puk89] models the increment in basal area and not the increment in dbh. Pukkala concludes that "The main reason for the good degree of determination is that the models were for basal area growth instead of diameter growth, and basal area growth correlated very closely with the diameter.". We could take $\mathrm{dbh}_t^\delta - \mathrm{dbh}_{t-1}^\delta$ as the response and again be interested in estimation of δ, which would probably be non trivial.

If we should answer the question "survival of the fattest?" in the title of this report in a short way, the answer must be yes. It is evident that it is the small trees and the trees with stronger neighbours that die first.

Acknowledgments

I would like to thank Antti Penttinen for his suggestion to use "energy marks" as a basis for the definition of competition indices and Anders Brix for many helpful discussions during the work.

References

[ABGK93] Per Kragh Andersen, Ørnulf Borgan, Richard D. Gill, and Niels Keiding, *Statistical models based on counting processes*, Springer Series in Statistics, Springer-Verlag, 1993.

[EW95] Jens Elberling and Ulrik Winther, *Noter til træmåling*, teaching notes, Institut for Økonomi, Skov og Landskab, Sektion for Skovbrug, 1995.

[FT94] Ludwig Fahrmeir and Gerhard Tutz, *Multivariate statistical modelling based on generalized linear models*, Springer Series in Statistics, Springer-Verlag, 1994.

[Heg74] Frank Hegyi, *A simulation model for managing Jack-pine stands*, Growth Models for Tree and Stand Simulation (Jören Fries, ed.), Royal College of Forestry, Sweden, Research Notes, no. 30, 1974, (Skogshögskolan, Rapporter och Uppsatser), pp. 74–90.

[Hei97] Juha Heikkinen, *Bayesian smoothing and step functions in the non-parametric estimation of curves and surfaces*, Phd thesis, University of Jyväskylä, Department of Statistics, 1997, Jyväskylä Studies in Computer Science, Economics and Statistics no. 40.

[Hen58] H. A. Henriksen, *Sitkagranens vækst og sundhedstilstand i Danmark*, Beretning nr. 191, vol. 24, Det Forstlige Forsøgsvæsen i Danmark, 1958, pp. 1–372.

[Hen81] H. A. Henriksen, *Træmåling*, 1981.

[Hen88] H. A. Henriksen, *Skoven og dens dyrkning*, Nyt Nordisk Forlag Arnold Busck, Copenhagen, 1988, 664 pp.

[HP95] Juha Heikkinen and Antti Penttinen, *Bayesian smoothing in estimation of the pair potential function of Gibbs point processes*, preprint of paper 2 in [Hei97], 21 October 1995, 1–22.

[Joh96] Vivian Kvist Johansen, *Statistical aspects of a spatial single-tree growth model for even-aged oak stands in Denmark*, Dynamic Growth Models for Danish Forest Tree Species Working Paper No. 10, Danish Forest and Landscape Research Institute, Hørsholm, 1996.

[Jør94] Bruno Bilde Jørgensen, *Metodik ved anlæg og måling af afdelingen for skovdrift's prøveflader*, note, FSL, 1994.

[McC80] Peter McCullagh, *Regression models for ordinal data*, Journal of the Royal Statistical Society, series B **42** (1980), no. 2, 109–142, with discussion.

[MN89] P. McCullagh and J. A. Nelder, *Generalized linear models*, second ed., Monographs on Statistics and Applied Probability, no. 37, Chapman & Hall, 1989.

[Nes96] Marks R. Nester, *An applied statistician's creed*, Applied Statistics **45** (1996), no. 4, 401–410.

[OSV97] D. G. Oreshkin, J. P. Skovsgaard, and J. K. Vanclay, *Estimating sapling vitality for Scots pine (Pinus sylvestris L.) in Russian Karelia*, Forest Ecology and Management (1997), 147–153.

[Pen97] Antti Penttinen, *Private communication*, 25–26 November 1997, initiated at "workshop: spatial statistics and GIS" and continued in e-mail afterwards.

[Pet96] Jørgen Holm Petersen, *Analyzing genetic and environmental influences on mortality*, Phd dissertation, University of Copenhagen, Department of Biostatistics, 1996.

[Phi94] Michael S. Philip, *Measuring trees and forests*, CAB International, 1994.

[PSH92] Antti Penttinen, Dietrich Stoyan, and Helena M. Henttonen, *Marked point processes in forest statistics*, Forest Science **38** (1992), no. 4, 806–824.

[Puk89] Timo Pukkala, *Methods to describe the competition process in a tree stand*, Scandinavian Journal of Forest Research **4** (1989), 187–202.

[RB85] David D. Reed and Haraold E. Burkhart, *Spatial autocorrelation of individual tree characteristics in loblolly pine stands*, Forest Science **31** (1985), no. 3, 575–587.

[RC94] Stephen L. Rathbun and Noel Cressie, *A space-time survival point process for a longleaf pine forest in Southern Georgia*, Journal of the American Statistical Association **89** (1994), no. 428, 1164–1174.

[SG91] Dietrich Stoyan and Paval Grabarnik, *Second-order characteristics for stochastic structures connected with Gibbs point processes*, Mathematische Nachrichten **151** (1991), 95–100.

[SGW] Hans T. Schreuder, Timothy G. Gregoire, and Geoffrey B. Wood, *Sampling methods for multiresource forest inventory*, Wiley.

[SJ95] Thomas H. Scheike and Tina Kold Jensen, *A discrete survival model with random effects: An application to time to pregnancy*, Research Report 95/11, Department of Biostatistics, University of Copenhagen, 1995.

[SJV98] J. P. Skovsgaard, V. K. Johansen, and Jerome K. Vanclay, *Accuracy and precision of two laser dendrometers*, Forestry **71** (1998), 131–139.

[SKM95] Dietrich Stoyan, Wilfrid S. Kendall, and Joseph Mecke, *Stochastic geometry and its applications*, second ed., Wiley, 1995.

[Sko97a] J. P. Skovsgaard, *Tyndingsfri drift af sitkagran. En analyse af bevoksningsstruktur og vedmasseproduktion i utyndede bevoksninger af sitkagran (*Picea sitchensis *(Bong.) Carr.) i Danmark*, Forskningsserien nr. 19, Forskningscentret for Skov & Landskab, Hørsholm, 1997, 525 pp.

[Sko97b] Jens Peter Skovsgaard, *Tyndingsfri drift af sitkagran*, Skoven **29** (1997), offprint.

[STSV95] Paula Soares, Margarida Tomé, J. P. Skovsgaard, and J. K. Vanclay, *Evaluating a growth model for forest management using continuous forest inventory data*, Forest Ecology and Management **71** (1995), 251–265.

[TB89] Margarida Tomé and Harold E. Burkhart, *Distance-dependent competition measures for predicting growth of individual trees*, Forest Science **35** (1989), no. 3, 816–831.

[TJSM+97] Mads Jeppe Tarp-Johansen, J. P. Skovsgaard, Søren Fl. Madsen, Vivian Kvist Johansen, and Ib Skovgaard, *Compatible stem taper and stem volume functions for oak (Quercus robur L and Q petraea (Matt) Liebl) in Denmark*, Annales des Sciences Forestières (1997), revised January 1997.

[Van94] Jerome K. Vanclay, *Modelling forest growth and yield; applications to mixed tropical forests*, CAB International, 1994.

[VS97] J. K. Vanclay and J. P. Skovsgaard, *Evaluating forest growth models*, Ecological Modelling **98** (1997), 1–12.

[WS96] Olga Wälder and Dietrich Stoyan, *On variograms in point process statistics*, Biometrical Journal **38** (1996), no. 8, 895–905.

A Description of Course

The PhD course in forest biometrics consists of three main parts:

- The reading and discussion of several forest related articles and parts of books. References are given below.
- Visits to several of the long term field experiments of The Danish Forest and Landscape Research Institute (FSL).
- This analysis of the data from experiment MBII.

The first two points were carried out mainly during a one month visit at FSL in December 1996 and January 1997, while the third point was carried out during the period August 1997 to February 1998.

During the course the following papers and books have been read in addition to the other papers in the reference list: [Hen81], [Hen88], [EW95], [Van94], [Phi94], [SGW], [Sko97b], [Sko97a], [Joh96], [TJSM$^+$97], [VS97], [STSV95], [Jør94], [OSV97], [SJV98], [Nes96], [RB85].

B The Data in Graphs

This appendix shows graphs of the data. Look in Section 2.3 for comments.

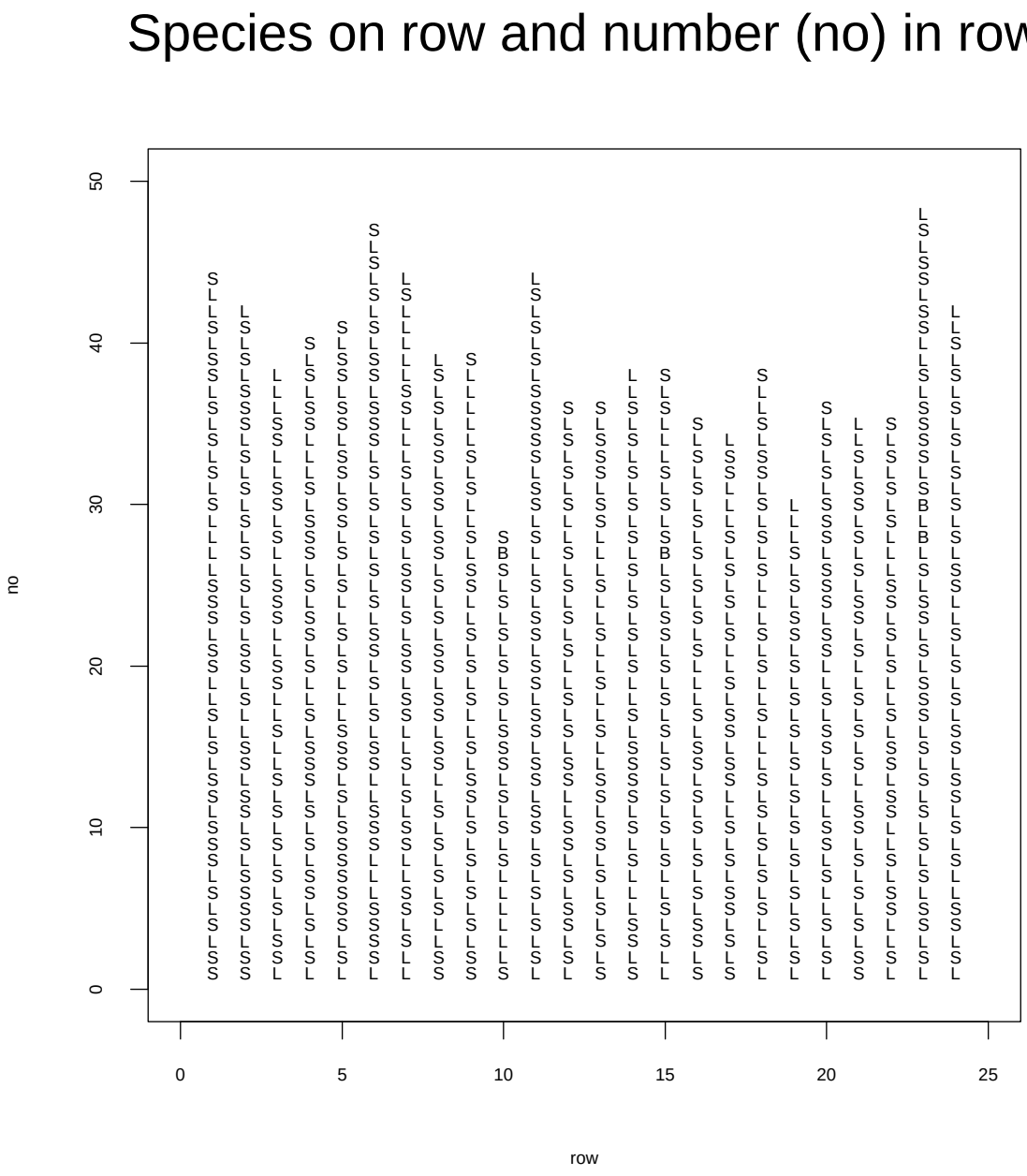

Figure 4: Positions in terms of row and number. L=larch, S=Sitka spruce, B=birch.

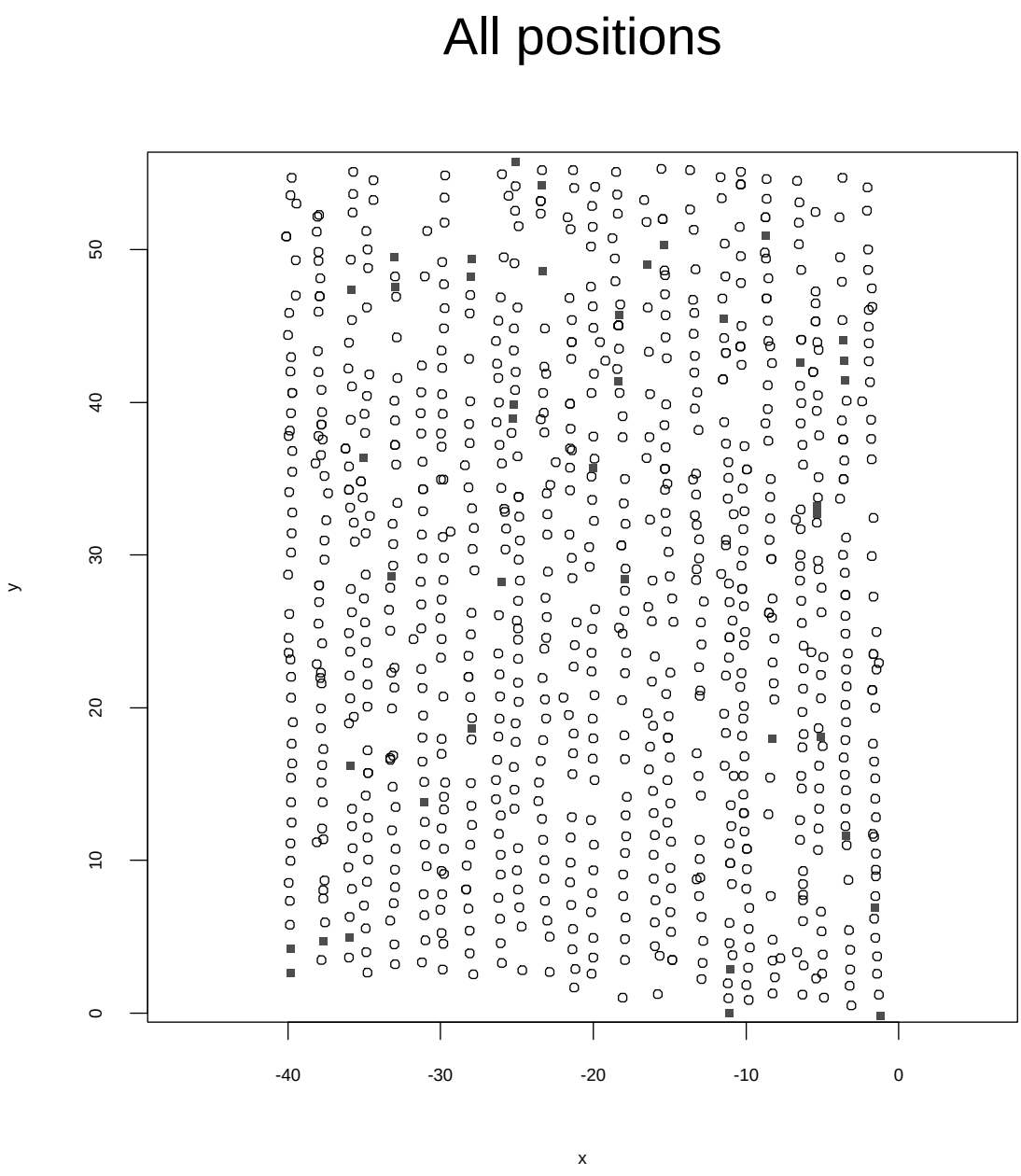

Figure 5: Positions of trees. The 41 defined positions are marked with a filled box.

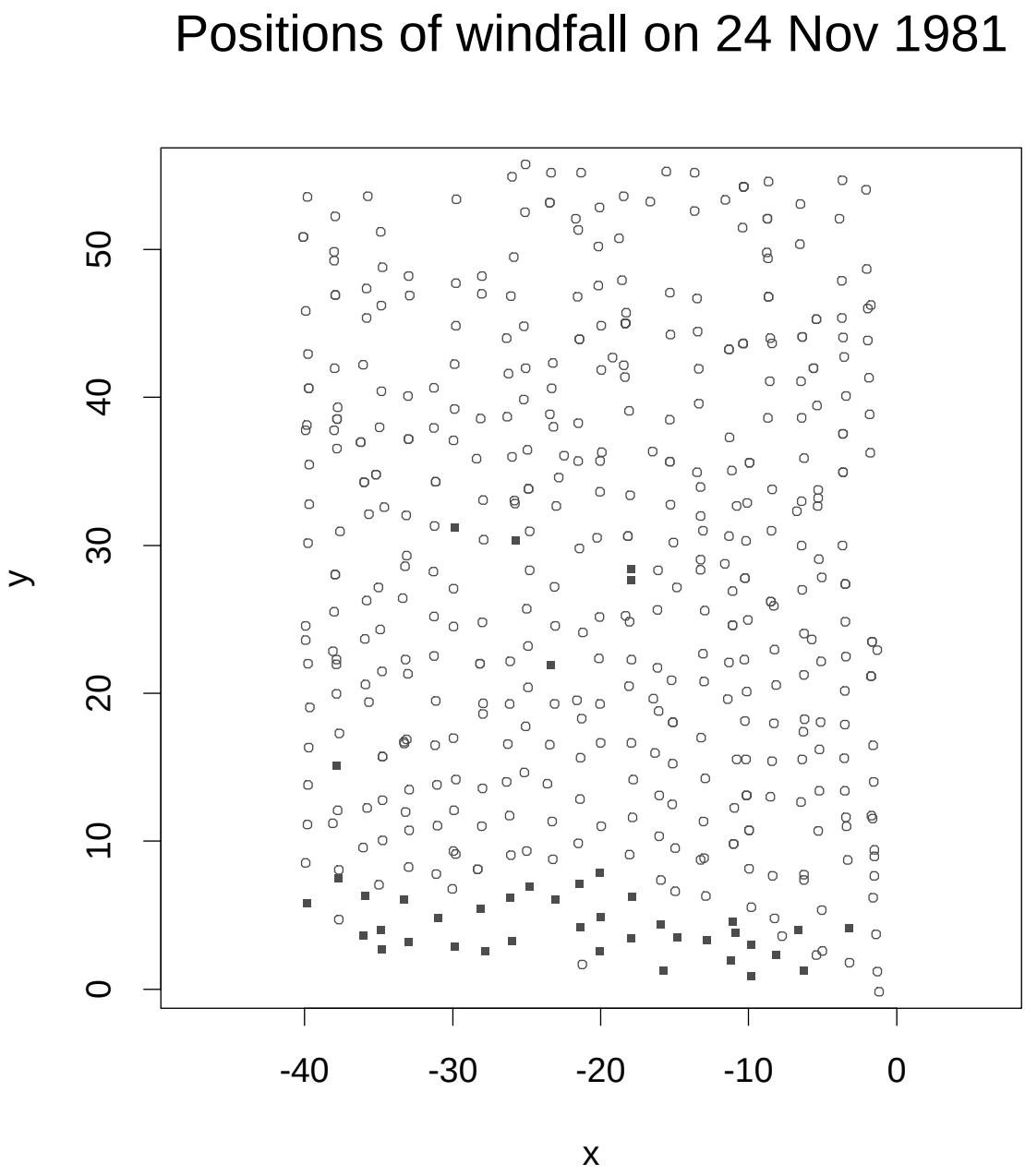

Figure 6: Positions of the fallen trees in the windfall are marked by filled boxes. Only the positions of Sitka spruce are plotted.

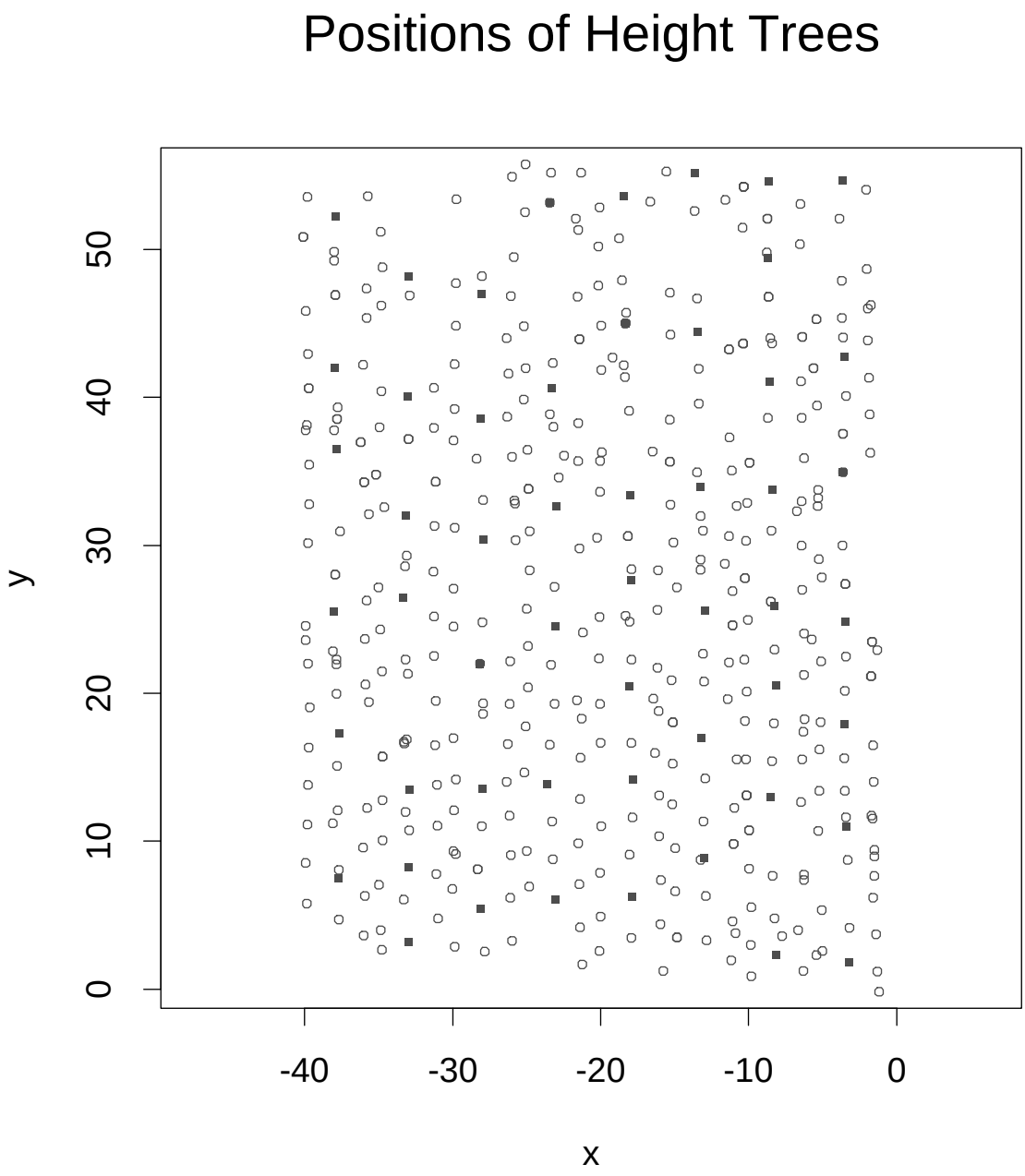

Figure 7: Positions of height trees are marked by filled boxes. Only the positions of Sitka spruce are plotted.

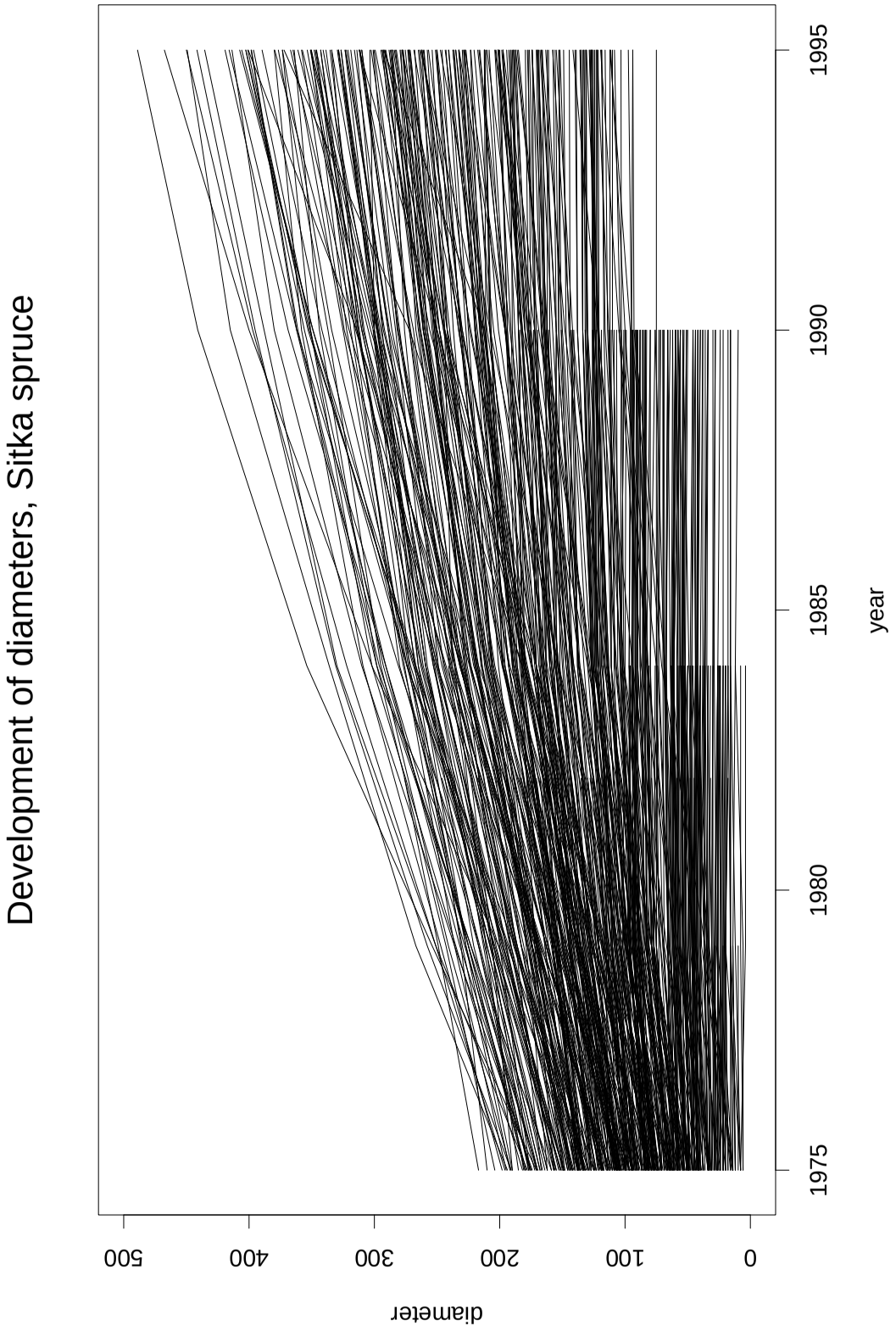

Figure 8: The longitudinal development of diameters.

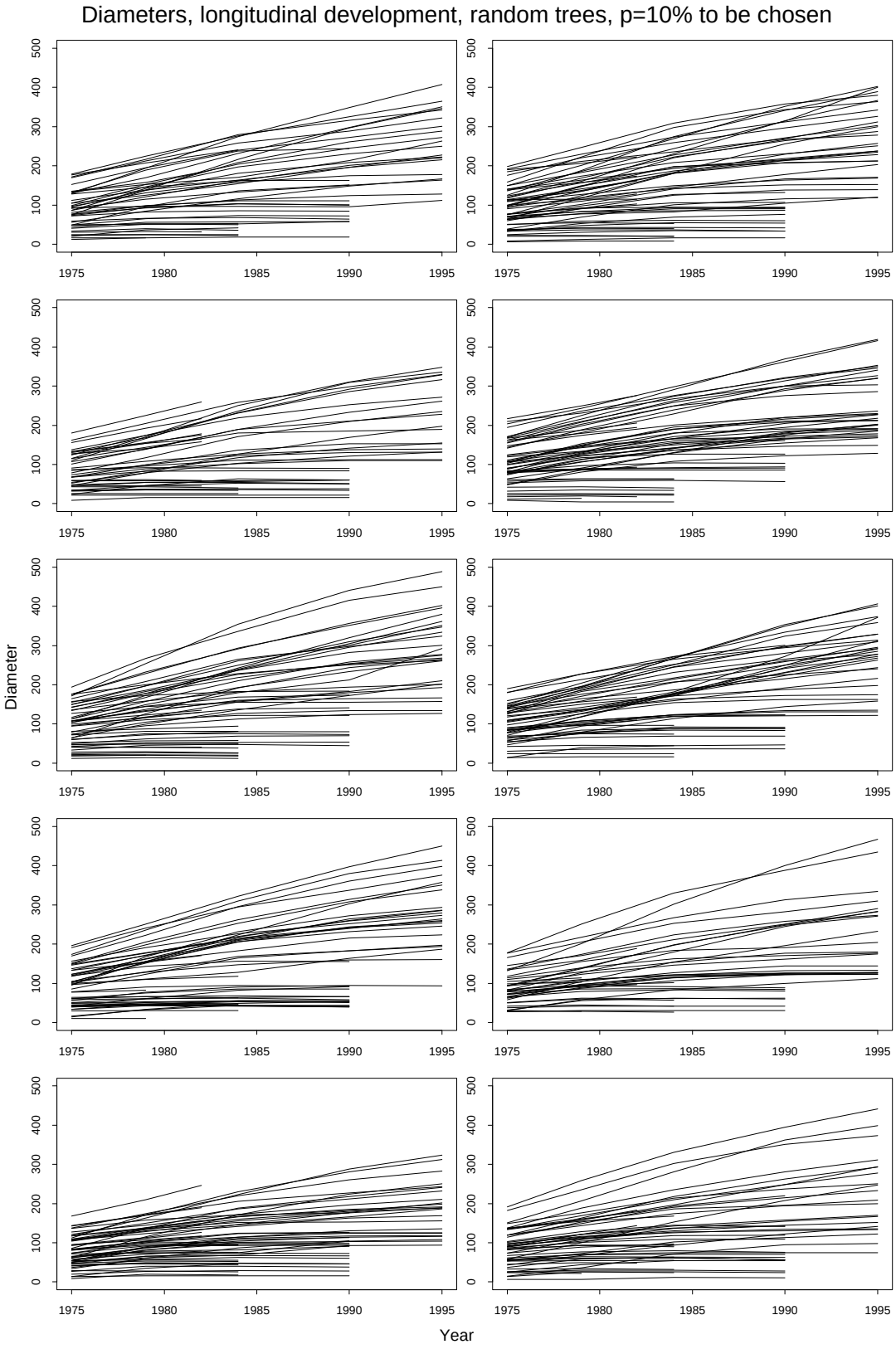

Figure 9: The longitudinal development of diameters; all trees are shown exactly once, and they are chosen at random for one of the ten plots.

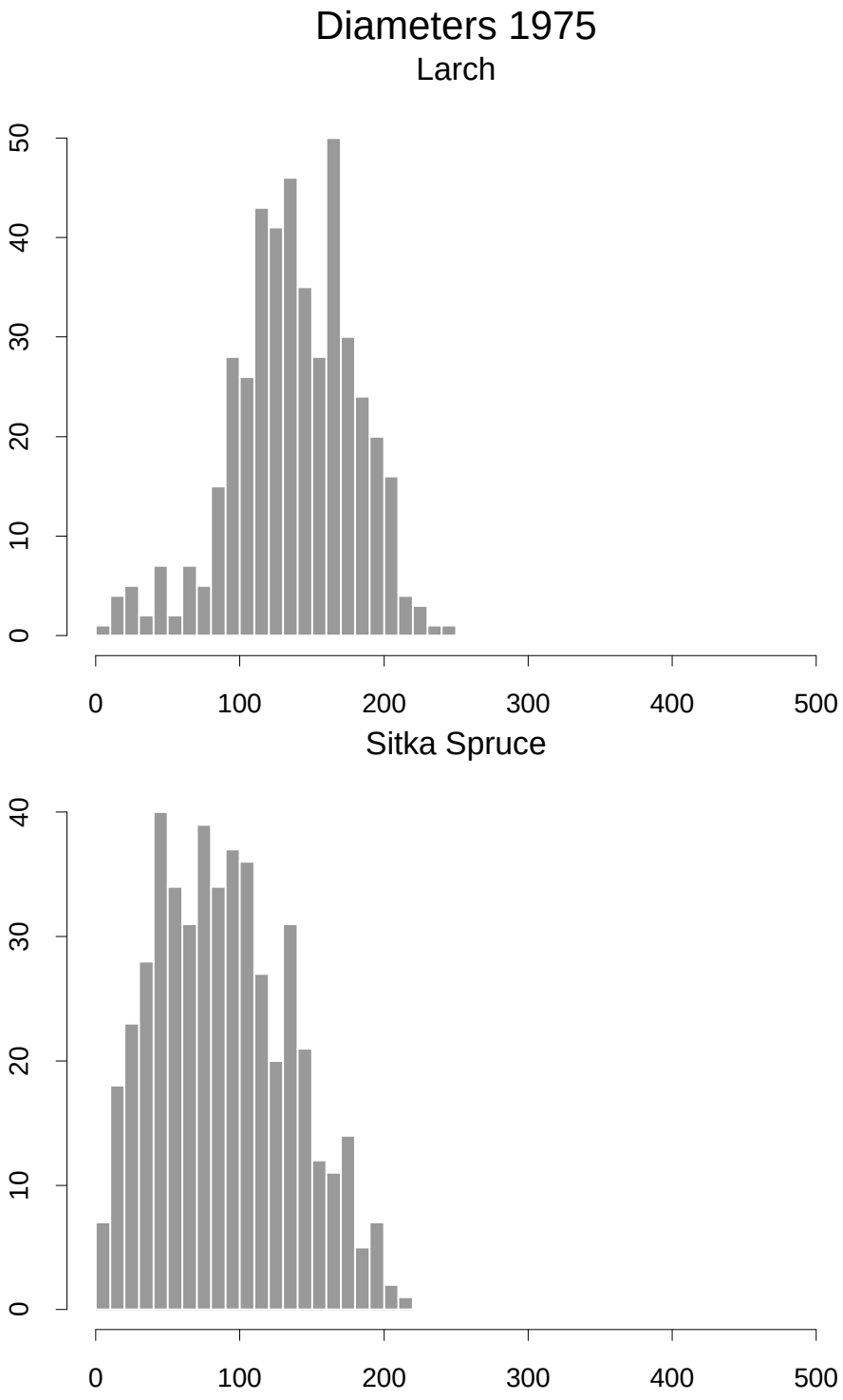

Figure 10: Diameters 1975, separated into species.

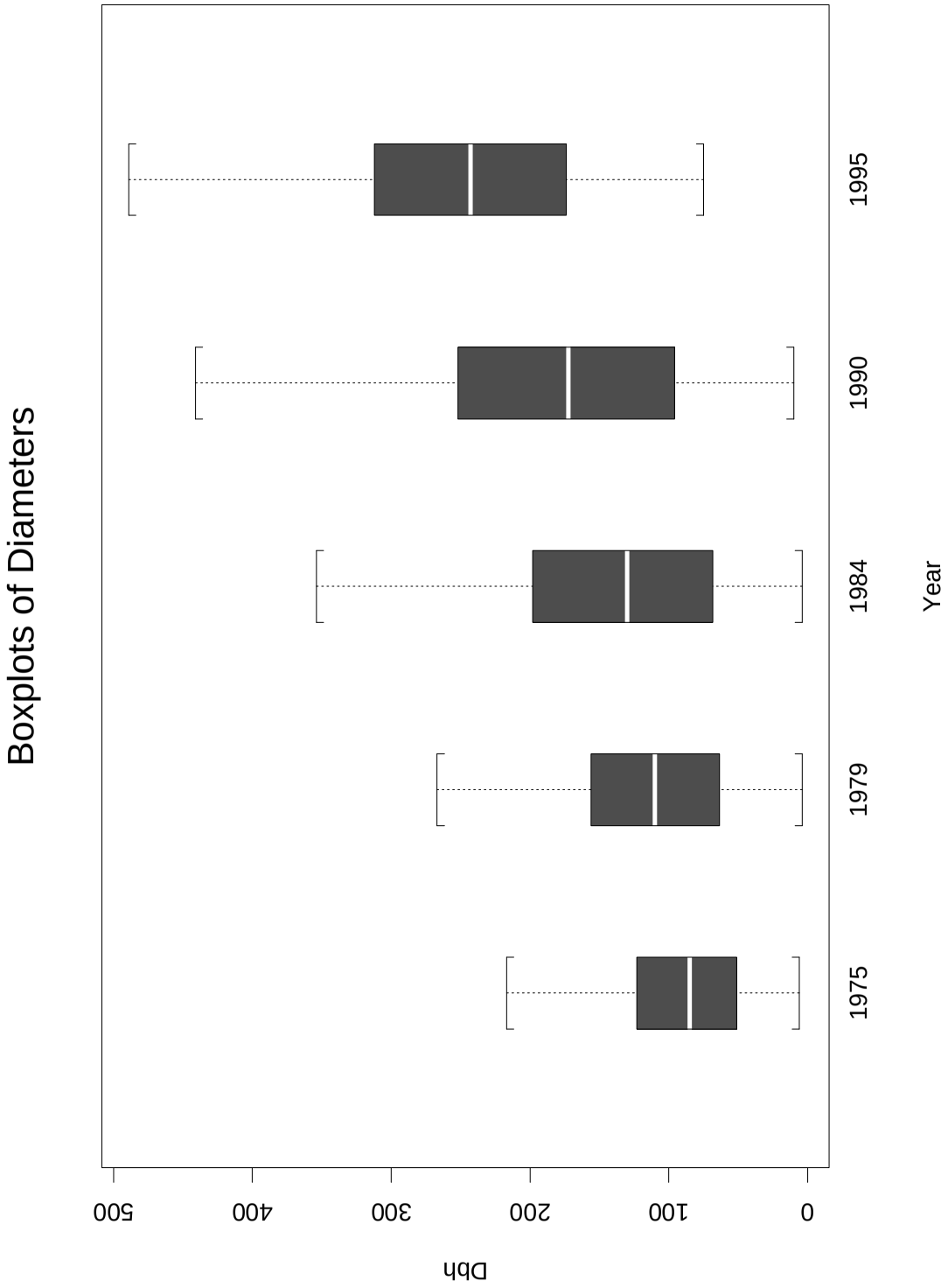

Figure 11: Boxplot of diameters. In 1975 only Sitka spruce is included.

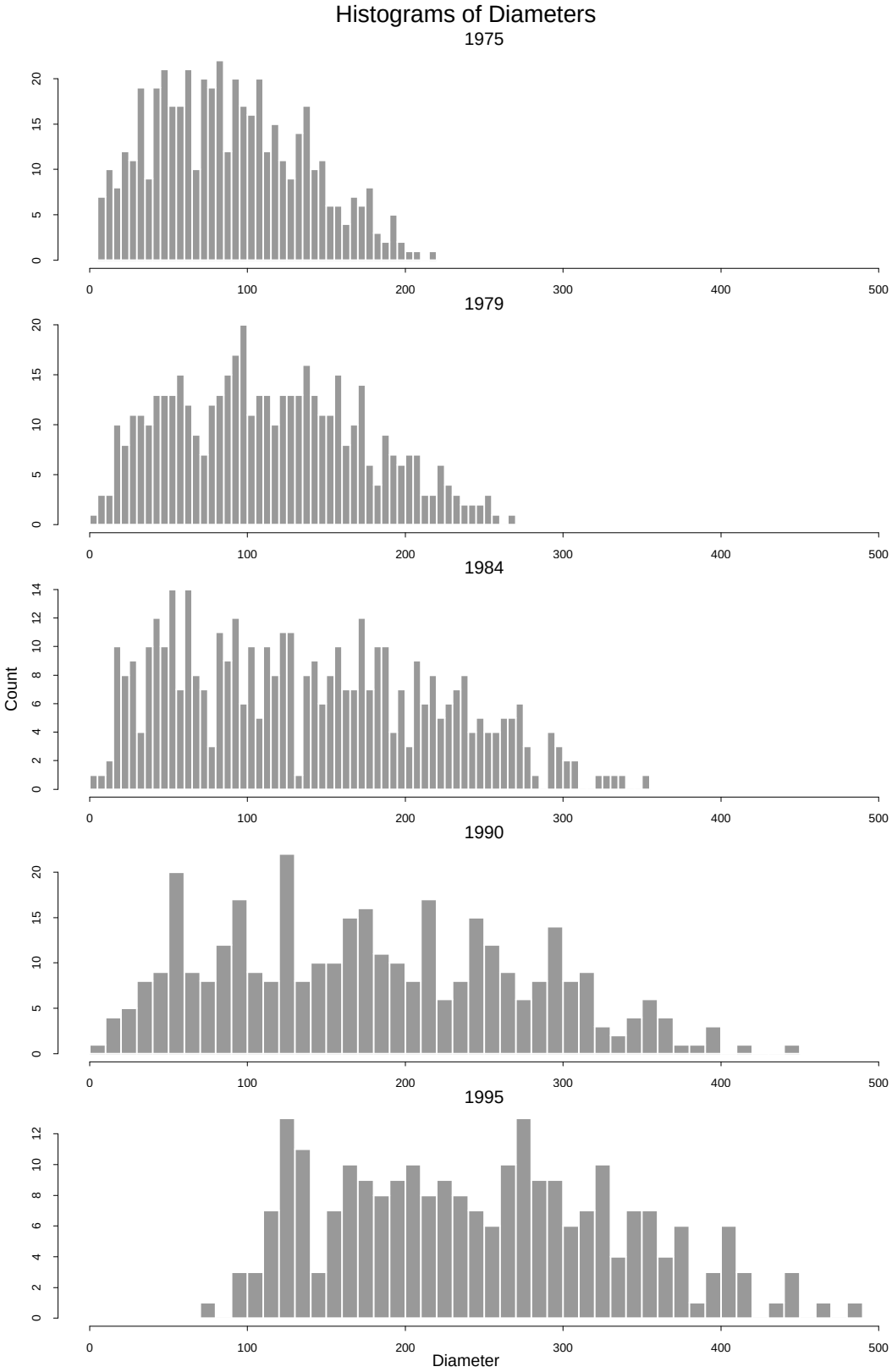

Figure 12: Histograms for the diameters. In 1975 only Sitka spruce is included.

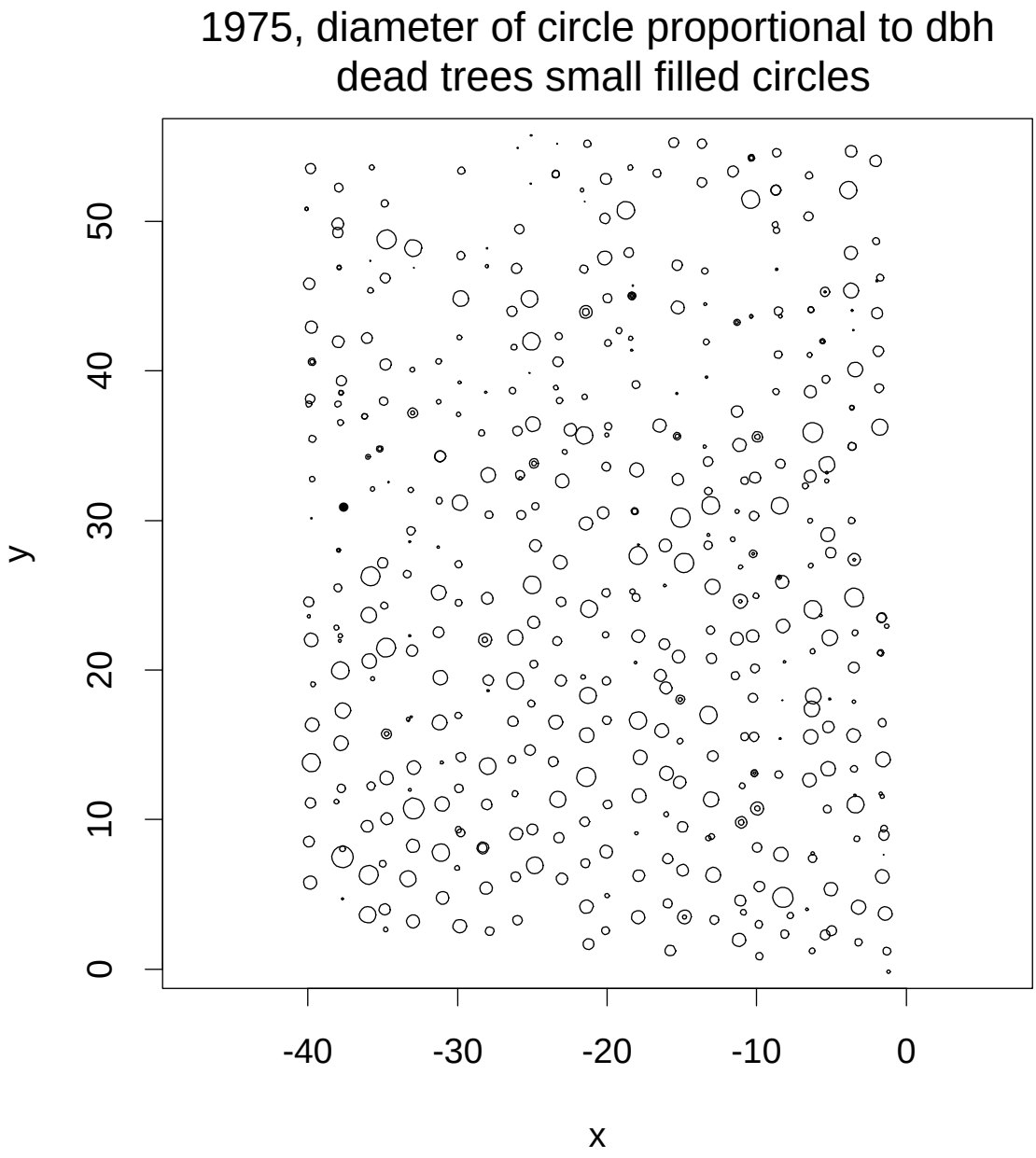

Figure 13: Diameters at positions, 1975. Dead trees are marked with dots of a fixed size. Only Sitka spruce.

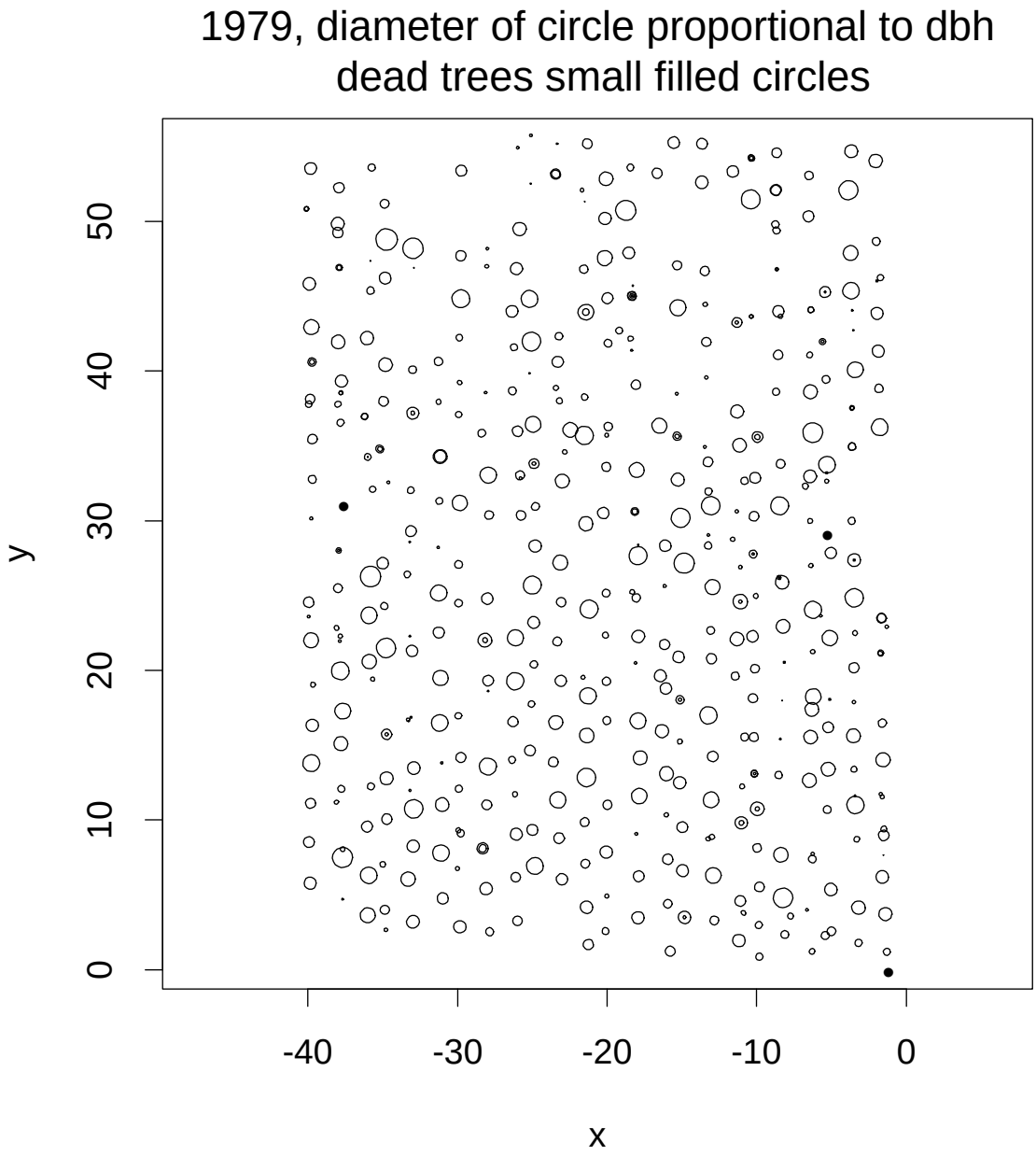

Figure 14: Diameters at positions, 1979. Dead trees are marked with dots of a fixed size.

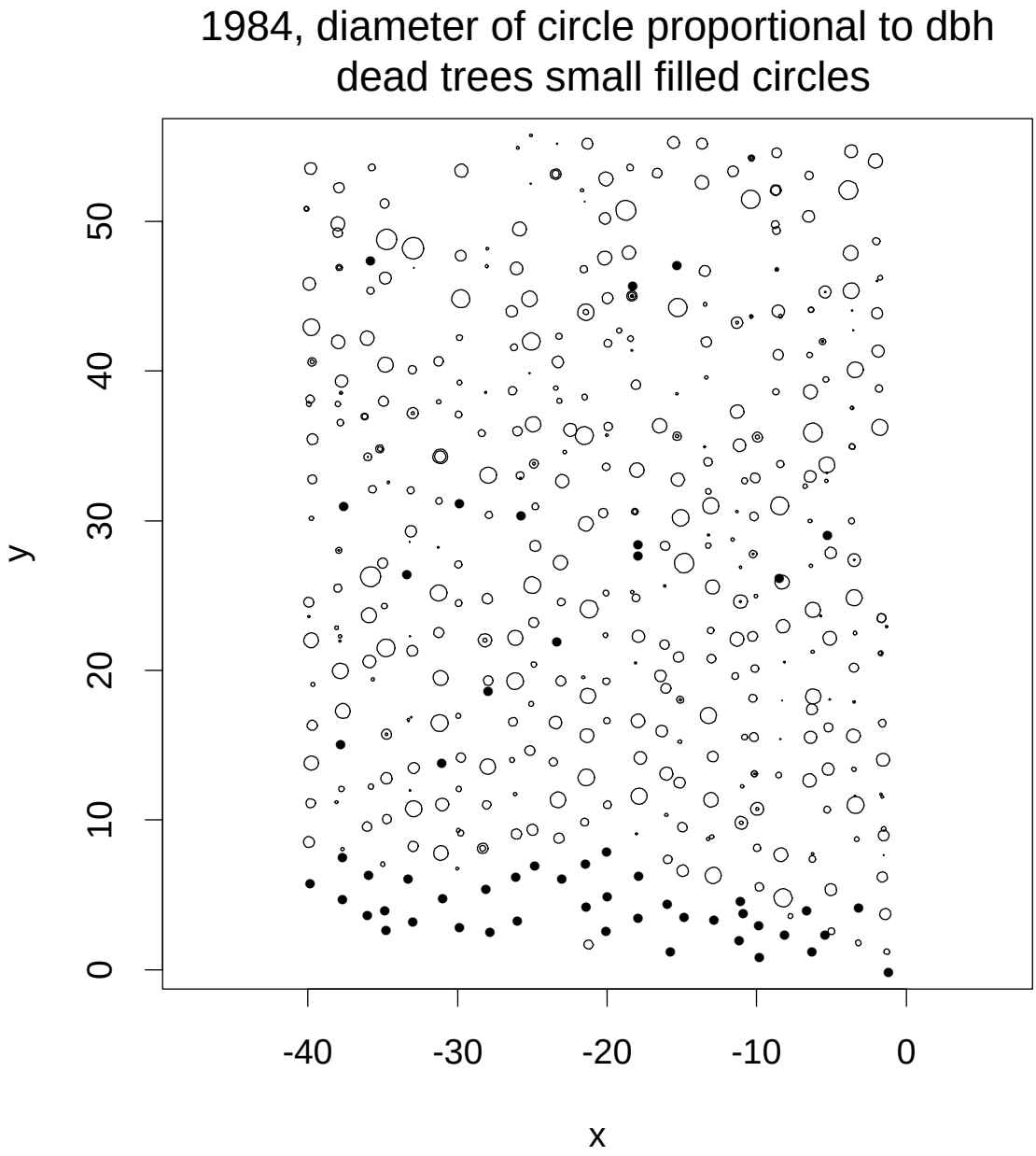

Figure 15: Diameters at positions, 1984. Dead trees are marked with dots of a fixed size.

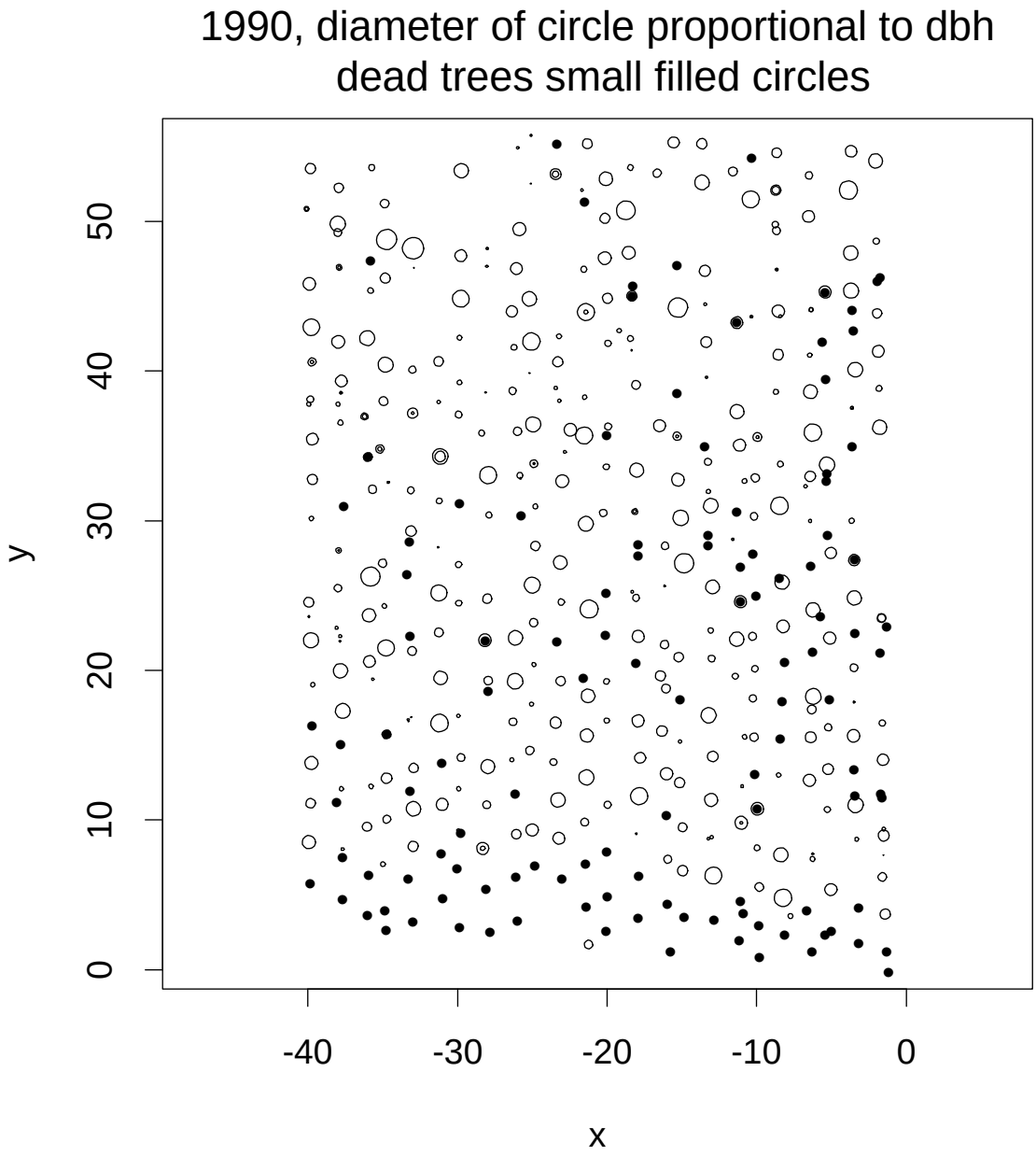

Figure 16: Diameters at positions, 1990. Dead trees are marked with dots of a fixed size.

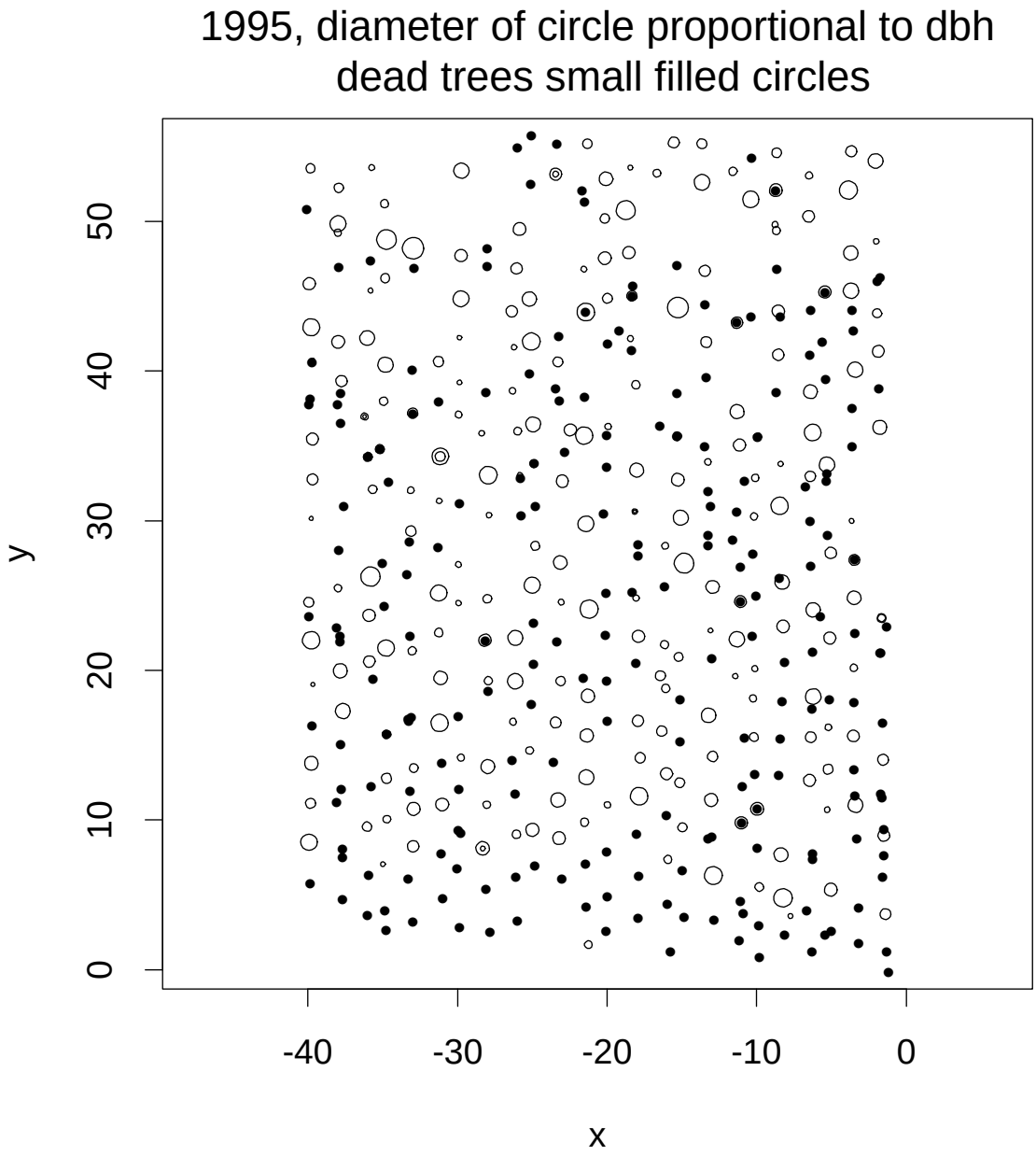

Figure 17: Diameters at positions, 1995. Dead trees are marked with dots of a fixed size.

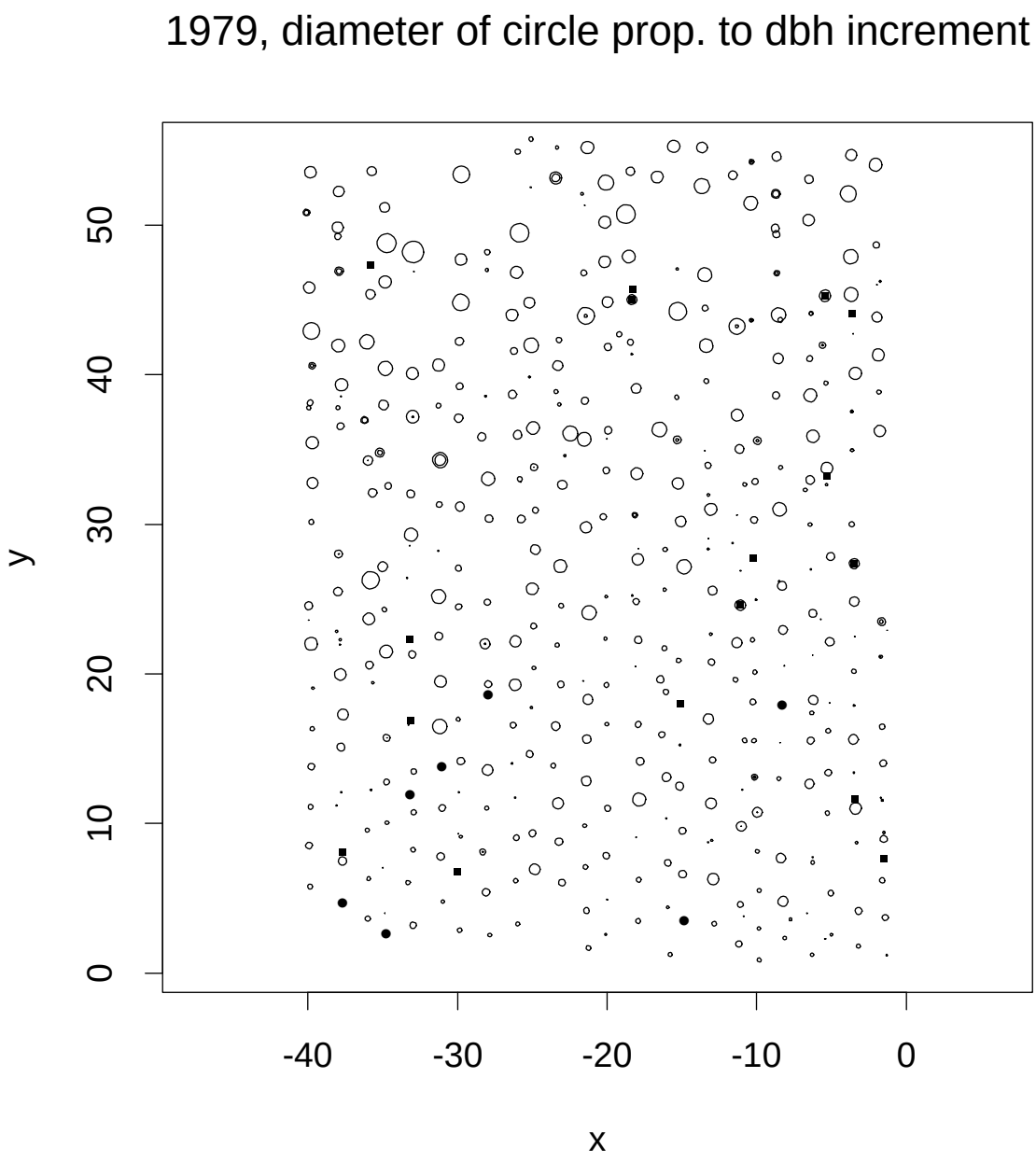

Figure 18: Map of increments between year 1979 and 1975. Small filled circles are negative increments, small filled squares are zero increments.

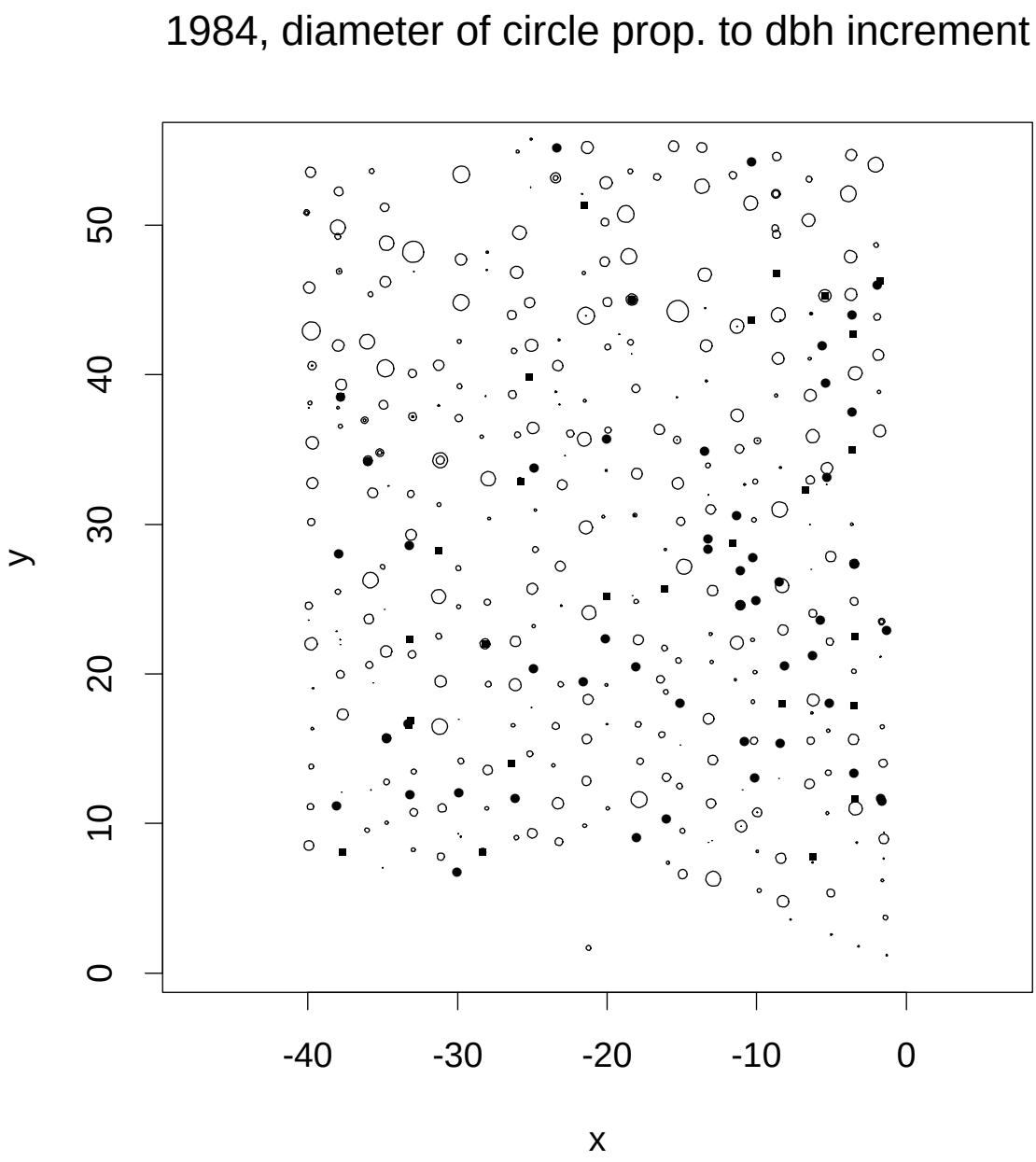

Figure 19: Map of increments between year 1984 and 1979. Small filled circles are negative increments, small filled squares are zero increments.

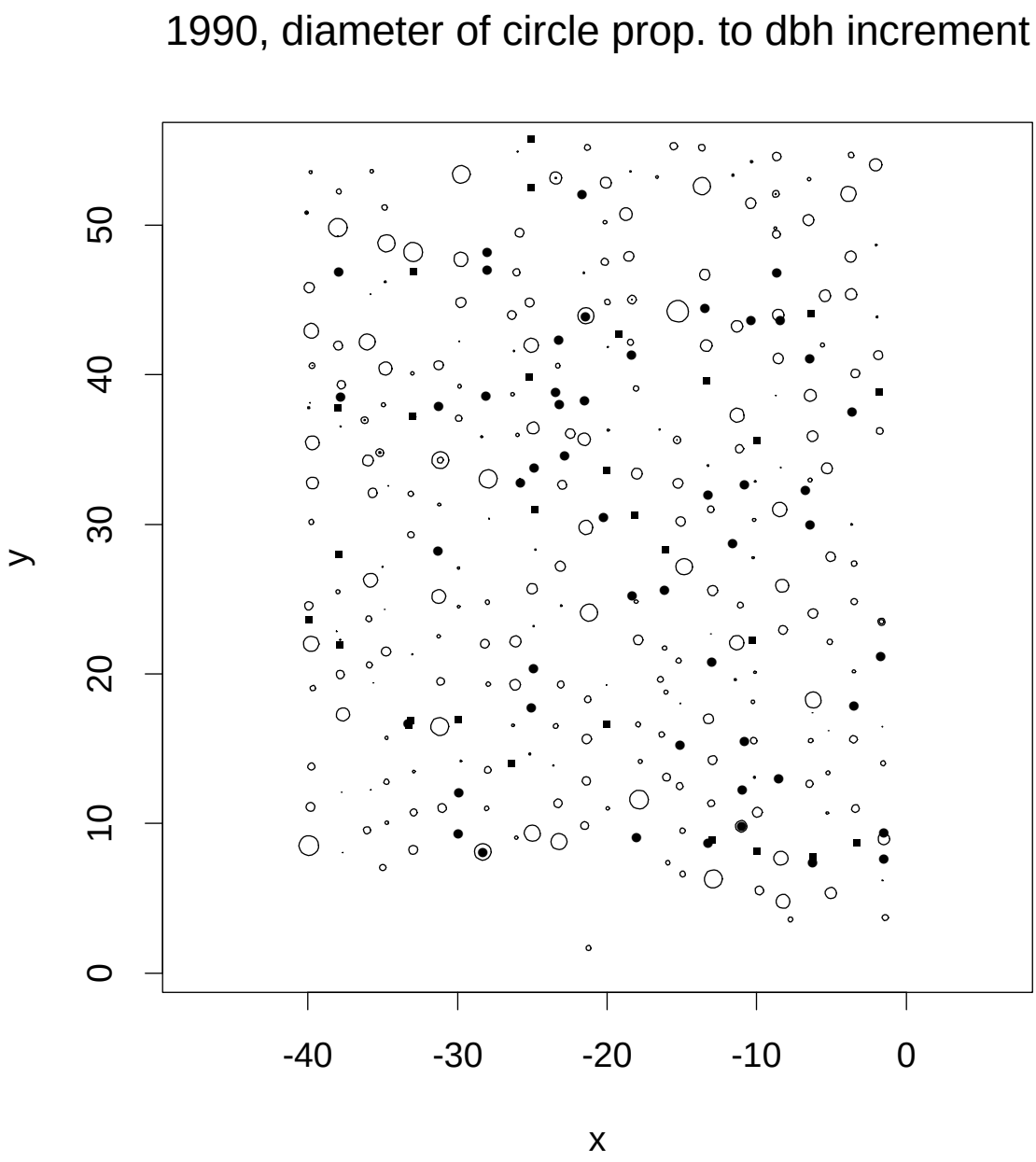

Figure 20: Map of increments between year 1990 and 1984. Small filled circles are negative increments, small filled squares are zero increments.

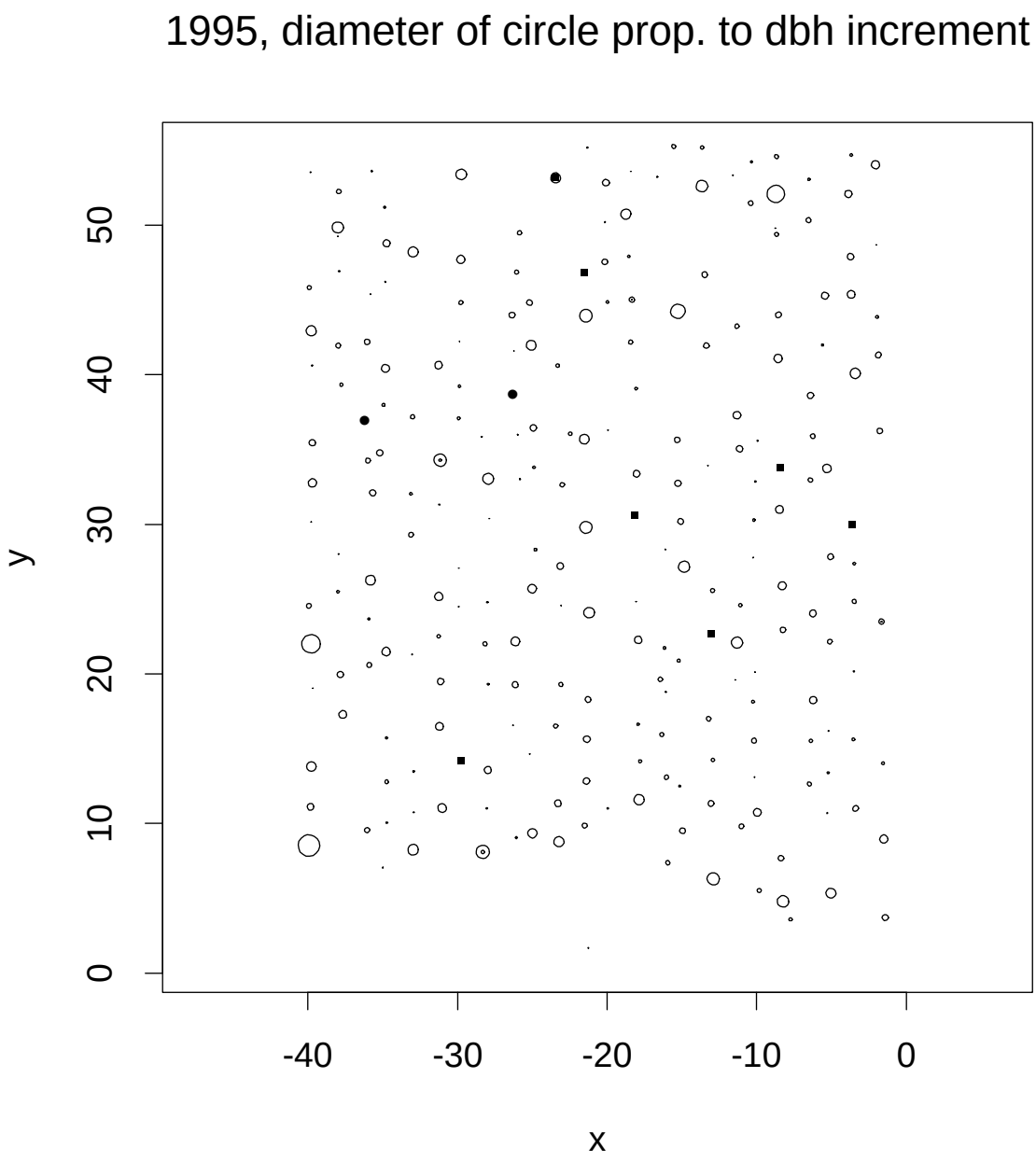

Figure 21: Map of increments between year 1995 and 1990. Small filled circles are negative increments, small filled squares are zero increments.

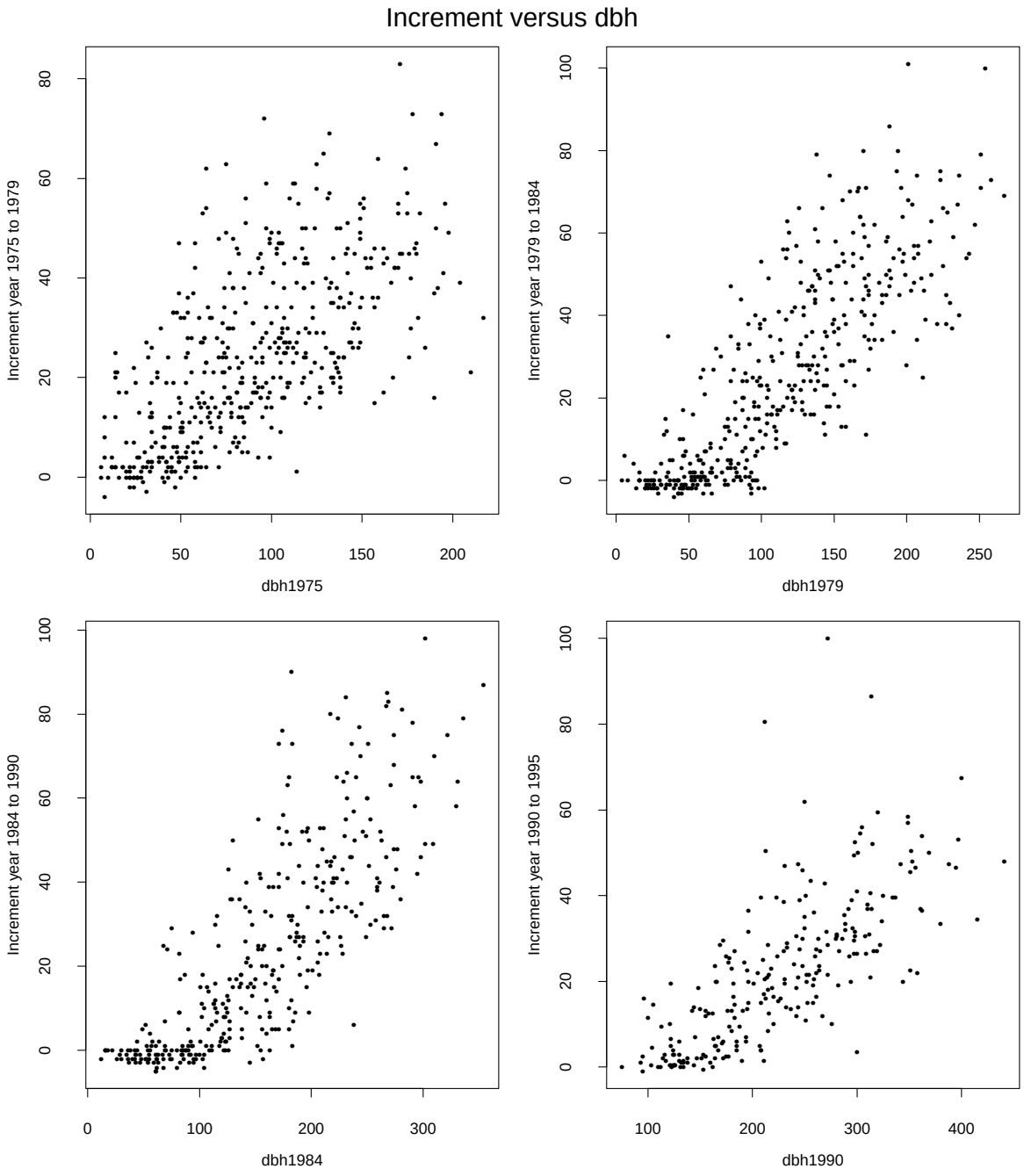

Figure 22: Plots of increment versus dbh.

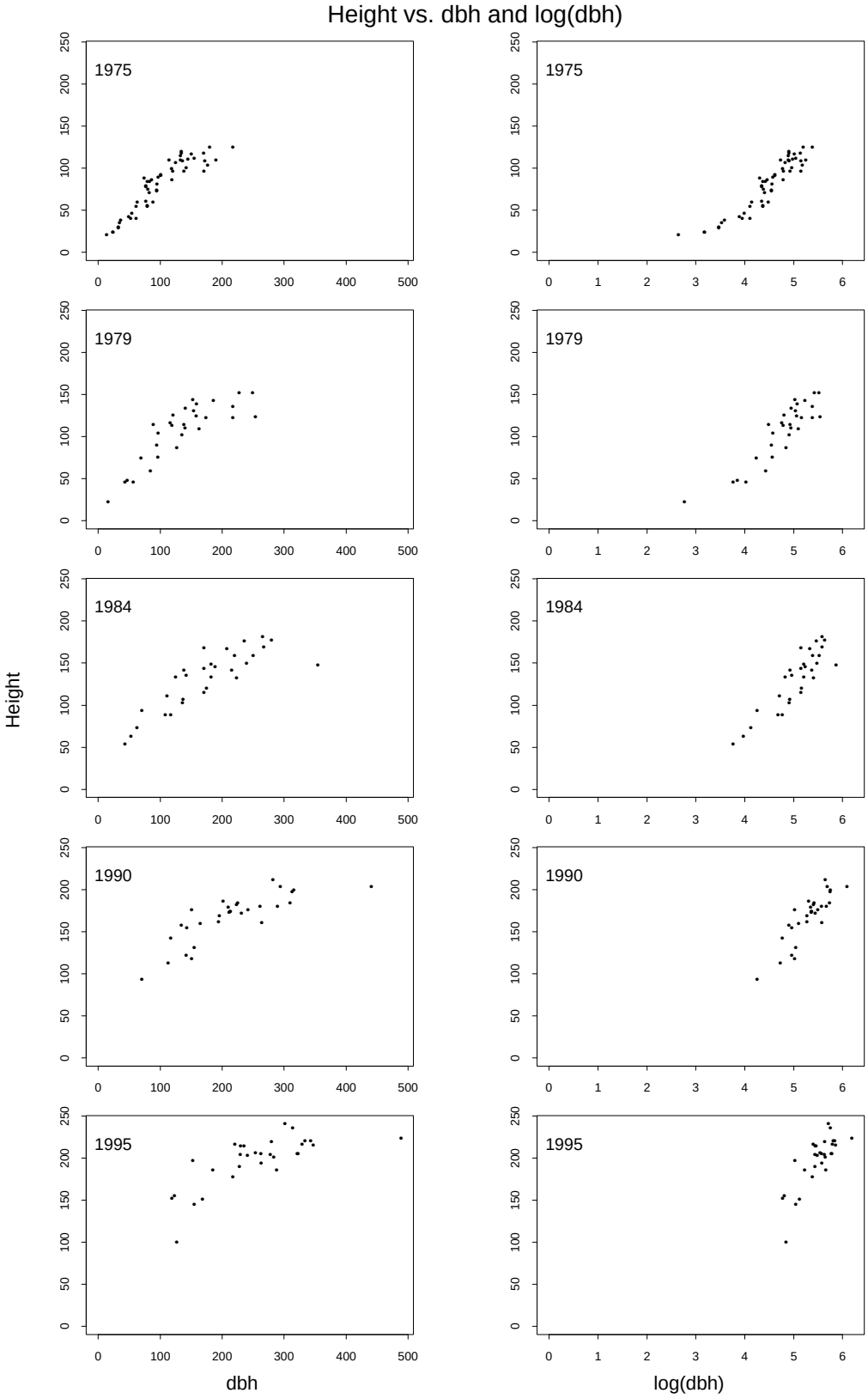

Figure 23: Plot of the heights versus the diameter at breast height, dbh, in the left column, and log(dbh) in the right column.