A Per-Plant Covariance Analysis Approach in the Study of the Effects of Missing Plants on Adjacent Plants in Field Experiments

Hector S. Paredes Herrera

Inst. of Statistics Mimeo Series # 1368

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued)

Page

LIST OF TABLES (Continued)

1. INTRODUCTION

Missing observations may occur in experiments conducted in several scientific areas of study. In particular, missing plants frequently occur in agricultural and forestry field experiments thus providing missing observations and affecting the interpretation of experimental results. This condition has motivated some researchers to undertake the investigation of problems relating to missing plants.

The causes of missing plants are numerous and of a diverse nature. Some of these are the choice of experimental material, seeding depth and rate, germination failure, transplanting damage, fertilizer injuries, mechanical injuries during cultural practices, attack by insects and diseases, space limitation, seasonal effects, storms, rainfall patterns, soil and air temperature, and solar radiation. Most field crop experiments are subject to the occurrence of missing plants, but the relative importance of these different causal factors varies from crop to crop.

When missing plants exist in an experimental field, they leave an empty space (gap) between plants introducing an environmental source of variability as the standing plants adjacent to these gaps are favored by having more available growing factors. Consequently, these standing plants are expected to respond differently from those plants with no adjacent gaps. The effects of missing plants on their neighboring standing plants is not uniform for all crops and experimental situations. It depends on some factors like variety, soil, season, climatic condition, plant spacing, size and shape of gaps, and the orientation of the gaps around a standing plant. Usually, some yield compensation by plants adjacent to gaps takes place. On the other hand, gaps sometimes may result in a lowering of yields of adjacent plants. The net effect upon plot yield is usually anticipated to be a decrease in relation to the yield of a plot having no missing plants. In some experiments, a high percentage of missing plants may render the data useless for analysis, especially if small plots are used.

There has been more of a tendency to ignore the effects of missing plants and to analyze plot yield data when the plants under investigation are small annuals such as cotton, peanuts, soybeans, etc. However, for high income tree or bush perennial crops where yield is measured by individual plant and where the number of plants per plot is limited, the effects of missing plants cannot be ignored. Moreover, the missing plants may have long-term effects in perennial crops because of the duration of experiments conducted on these crops. In these situations, missing plants not only destroy the balance property of the data (thus rendering analysis difficult), but they also may give biased estimates of the treatment effects and biased tests of significance. The precision of the experiments may also be decreased through inflation of the experimental error.

Some agronomic practices can be performed in order to prevent the occurrence of missing plants. Good nursery and establishment techniques should be conducted. A general practice for having a perfect stand of small plants is to plant thick and then to thin to a uniform rate. An alternative which has been used for both small and large plants is to replant under the same experimental conditions of the standing plants. The replanted plants do provide competition to adjacent plants. Also, in many cases it is difficult to reproduce the same experimental conditions for the replants as would have existed if the original plants remained standing. Replanting is sometimes impractical and it may be costly.

Several researchers have developed procedures for adjusting plot yields based on an estimated amount of yield compensation on plants adjacent to missing plants. Some of the procedures that have been developed involve judgment based on the researcher's own experience (Jones and Collins, 1959), whereas most of them involve formulas for adjusting plot yields derived from experimental results (Stewart, 1919; Pope, 1947; Crews and Jones, 1962; Bush and Ergun, 1973; Gupton and Archer, 1973). These formulas have been tested for particular crops (cotton, potato, tobacco, spring wheat) under various cultural situations, but it would not seem that they are entirely reliable for general application. Some authors working with corn, potato, and rice did not recommend the use of adjustment formulas because the effects of missing plants vary with some agronomical and/or environmental factors (Kiesselback, 1923; Myers and Perry, 1923; Giesbrecht, 1961; Hindi, 1962; Gomez and De Datta, 1972). In lieu of adjustment, the following have been recommended: (1) harvesting all plants, including the ones adjacent to gaps, and basing the plot yields on area harvested (assuming total compensation) or on per plant harvested (assuming no compensation); or (2) discarding plants adjacent to gaps and basing plot yields (assuming some compensation) on plants completely surrounded by standing plants.

Allan and Wishart (1930) and Yates (1933) were among the first to develop missing plot formulas for the estimation of missing plot values in the analysis of incomplete data. The analysis of covariance has been an important statistical method to correct plot yields on the basis of stand irregularities ever since Fisher (1932) proposed this technique. Missing plot formulas and the analysis of covariance have been widely accepted and used by researchers as standard missing plot techniques.

In this thesis, a new approach has been studied for elucidating the nature of the missing plant effects and for adjusting for missing plants in agricultural and forestry experiments. The approach involves the method of the analysis of covariance where the covariates are measured on a per-plant basis. The reason for using a per-plant basis is that a plant adjacent to a missing plant is the plant directly affected by the missing plant. This effect on plants adjacent to gaps is a bias in the response. The response may be positively or negatively biased if the adjacent plants overcompensate or undercompensate, respectively, for the loss of response due to the missing plants. An unbiased response per unit area could be obtained if the adjacent plants do not compensate for the loss of yield; however, this is not the usual case.

The objectives of the study were to:

1. Derive a general covariance procedure for determining the effects of missing plants on adjacent standing plants. This includes development of a set of principles for deciding which covariates should be included in the covariance model.
2. Fit covariance models of the type developed in (1) to several sets of crop and forestry data using covariates specific to the particular set of data.
3. Develop a procedure to adjust treatment means for the effects of the covariates which were determined to be the most relevant for estimating the missing plant effects in each data set.

4. Estimate the total increase or decrease of the response variable measurements due to the missing plant effects for each treatment and each entire experiment studied.
5. Analyze the data sets after the estimated missing plant effects have been removed from the treatment effects.
6. Compare the results obtained in (5) with analysis of covariance which uses number of standing plants per plot as the covariate.

2. REVIEW OF LITERATURE

Detailed information concerning the effects of missing plants on adjacent standing plants is very scarce in the literature. The major emphasis has been upon the net effect of the missing plants upon the yield of the plot.

Most studies have been conducted on annual crops. Pope (1947), working with cotton, recommended replanting the entire row or field if many long skips are present. He found a significant linear reduction in yield due to skips of more than three feet in single-row plots for which he presented adjustment formulas. For multiple-row plots, no adjustment was made because skips in interior rows are largely compensated for by increased production of end plants in the row containing the skip and by lateral compensation of the adjacent rows. Moursi (1956), using different missing plant patterns as treatments, found that a 12% per-plot reduction in stand did not significantly reduce the yield of dasheen; moreover, the yield compensation decreased as the number of missing plants increased from one to three. Gomez and De Datta (1972) studied the effect of missing hills on surrounding hills in experiments on rice using seven missing hill patterns as treatments. They reported that the surrounding plants gave greater yields but that a correction factor was not applicable because the yield compensation varies with variety, plant spacing, fertility level, crop season, and the number and distribution of missing hills. They recommended harvesting only fully competitive plants. Bush and Ergun (1973) concluded that the yield of spring wheat decreased significantly when the missing plant area was more than 12.5% of the plot area. They presented formulas for adjusting the yields per plot but they reported that overadjusted plot yields may be obtained. Teigen and Vorst (1975) reported that the stand reduction levels studied (25% and 50%) reduced seed yield of soybean, but that remaining plants compensated for those removed by producing more pods per plant.

Several research workers have studied the effects of missing plants on yield in experimental plots involving corn. Kiesselbach (1918) suggested that only the same number of fully competitive hills for each plot should be harvested for more precise comparisons. He indicated that a standing plant compensated for the loss in yield to the extent of 15%, 9%, and 2% when it was adjacent to a hill with 3, 2, and 1 missing plants, respectively, in relation to a normal hill surrounded by normal hills. He did not comment on the effects of missing diagonal hills. Later (1923) he found similar yield compensation results (14%, 7%, and 2%) for the vertical and horizontal hills. Additionally, he mentioned the possibility of compensation due to missing diagonal hills but he did elaborate. He further added that rather exaggerated yields would often be obtained by harvesting all full hills irrespective of the adjacent stands if there exists many hills with missing plants. Moreover, he indicated that one cannot rely on a general formula for correcting yields on a percent stand basis due to great variability from year to year in the relation between stand and yield. Therefore, he recommended to (1) plant thick and then thin to a uniform rate, and (2) base the yield upon normal hills surrounded by normal hills. Brewbaker and Immer (1931) agreed with Kiesselbach (1923) in saying that the response due to missing plants may vary considerably from year to year and that there may be a correlation between the response obtained and the yield. They also pointed out that missing plants may inflate the experimental error. They concluded that it is doubtful if competition exists beyond the three-feet spacing. Giesbrecht (1961) reported that hills directly adjacent to missing hills yielded significantly more than normal hills whereas hills diagonally adjacent to missing hills were not significantly different from normal hills. He did not recommend an adjustment factor but instead presented two methods to remove the effect of the missing hills. These methods are (1) "spiking in" (the seeding of some hills at a later date than the bulk of the test) with corn or sunflower, and (2) filling in with corn transplants. Hindi (1962) reported that the size and shape of skips, type of stand, amount of nitrogen applied, hybrids used, and climatic conditions may influence the effect of skips on the adjoining hills; therefore, an adjustment formula is not advisable because it would be misleading in its conclusions. He also indicated that the effect of skip was significant at the 1% level of significance and that the orientation of the hills around a skip is very important in the compensation for the loss in yield. Hills located to the north compensated 7.6-9.5% more than those located to the south and hills located to the west compensated 5.0-6.0% more than those located to the east. In the presence of missing hills he suggested that plot yield be estimated by the "normal plant method," i.e., to discard hills affected by missing hills, to find an average yield per hill and then to multiply by the number of hills that there were supposed to be at planting time.

Experiments on the effects of missing plants on potato growth have been conducted by Fitch and Bennett (1910), Stewart (1919), and Myers and Perry (1923). Fitch and Bennett and Stewart reported a 50%

compensation by the hills adjacent to missing hills. Myers and Perry reported that hills adjacent to single hill skips compensated as much as 46.4% and 53.8% in two different years, and with triple hill skips, the compensation was 43.8%.

Tobacco experiments have also been affected by the presence of missing plants. Jones and Collins (1959) developed through experience an adjustment procedure based on the assumption that increased growth of each plant adjacent to a missing plant results in a yield compensation equivalent to that of one-fourth of a plant. Crews and Jones (1962) investigated eight different patterns of missing plants at four stages of growth. The effects of both patterns and stages of growth were significant. They estimated the amount of compensation by adjacent plants and derived an adjustment formula for plot yields on the basis of stand. They indicated that the procedure should provide a good estimate of compensation under average environmental or seasonal conditions. The relative precision obtained by their procedure was greater than that for the procedures of Jones and Collins, than that using the covariance analysis of plot yields on stand, and than that in which no adjustment was used. Gupton and Archer (1973) developed an adjustment procedure similar to that of Crews and Jones using data from an experiment having five patterns of missing plants at four stages of growth as treatments. The effects of both patterns and stages were significant. The relative efficiency of their procedure was greater than that without adjustment, than that of Crews and Jones, and than that of covariance analysis of plot yields on stand. They pointed out that the effects of missing plants depend on variety, soil, fertilizer, and harvesting method.

Some papers have also been written on results from experiments on perennial plants. Pearce (1953) advised the practice of replanting. He suggested no adjustment of plot yields if the missing plants are due to the effect of some treatments. The data analysis would then be carried out on plot totals including the replant yields as well as the data on the original plants. Where adjustment is necessary, he recommended the use of the analysis of covariance on a plot basis with the number of gaps or replants as covariates in the analysis of covariance of plot totals or mean yield per surviving plant for each plot, ignoring replants. He also mentioned that the effects of missing plants upon adjacent plants depend largely upon the species. In a later paper (1955), he reported that the presence of missing plants has to be taken into account in deciding about plot size. In case there are missing plants in the experiment, he decided as follows: (1) for one-plant plots: use established methods of missing plot techniques; (2) for four-plant plots: discard plot data if there are missing plants or use approximate methods of analysis; and (3) for six-plant (or more) plots: work with the mean yield per plant for each plot and adjust by the method of covariance for the loss of competition if necessary. In the case of four-plant plots, the inaccuracy of the results on a plant basis can be real serious, he added.

In oil palm experiments, Haines and Benzian (1956) reported that young replants among older plants showed a growth which was too slow and which increased the error per plant. They did not present any adjustment procedure. Chapas (1961) mentioned that the loss of two plants in a plot of four would render the plot useless. For a few losses in large plots, he recommended the use of the analysis of covariance with the number of standing plants per plot as a covariate. A form of weighted adjustment on standing plants for effects due to missing plants is conducted at the Santo Domingo Experiment Station of INIAP, Santo Domingo, Ecuador.[1] The weights depend on the number of missing plants adjacent to a standing plant, and the planting pattern (square or triangular). The yield of a standing plant next to one or more missing plants is multiplied by an adjustment factor which is designed to reduce this yield to what it would have been had there been no missing plants. Their adjustment multipliers for a square planting are 1.00, 0.86, 0.79, 0.76, and 0.72 for 0, 1, 2, 3, and 4 adjacent missing plants, respectively. Those for a triangular planting are 1.00, 0.87, 0.77, 0.70, 0.63, and 0.58 for 0, 1, 2, 3, 4, and 5 adjacent missing plants, respectively. It is not clear what procedure was used for derivation of the adjustment multipliers.

Kowal (1959) replanted annually for the missing plants that occurred during the eight years of establishment of cacao in experiments under Nigerian conditions.

Abeywardena (1964) reported that missing plants introduced high variability in coconut yields because the gaps allowed some plants to have more growth area. However, he did not present procedures to adjust for the missing plots.

Forest research workers have also undertaken studies dealing with missing plants. In the application of field plot techniques for studies on plot size and shape, Wright and Freeland (1960) used the plot mean based on the number of survivors with the restriction that plots were

---

[1] Personal communication, Dr. L. A. Nelson, Department of Statistics, N. C. State University at Raleigh.

discarded if more than 50% of the plants were missing. They indicated that a 50% mortality might invalidate an experiment with one-plant plots. Their recommendations on how to deal with the missing plant problem were to (1) keep mortality low by using the best nursery and establishment techniques available; (2) replant for first-year mortality if the difference between replants and the original planting does not remain very long; (3) use standard missing plot techniques to estimate missing plot values, especially if total mortality is less than 5%; and (4) eliminate varieties or replicates with high mortality from the analysis and then make estimates for some missing plot values remaining in the experimental data. Woessner (1965), experimenting on loblolly pine, found plant survival ranging from 70-96%. He made no adjustment for missing plants nor did he apply missing plot techniques but he used plot mean values for the analysis of his data. He did this based on results given by Blake (1959, cited by Woessner) and Wright and Freeland (1960).

Some statistical textbooks by Scheffé (1959), Steel and Torrie (1960), Cochran and Cox (1966), Snedecor and Cochran (1968), and Searle (1971) present methods based on the least squares estimation procedure to deal with data having missing observations. These methods are applied to the existing plot data to perform an exact data analysis or to compute values by missing plot formulas. The computed values are then inserted in place of the missing ones to have a complete set of data on which analysis of variance, with some modifications, can be performed.

Another method which has been presented to adjust plot yields on stand basis is analysis of covariance (Mahoney and Beaten, 1953; Steel and Federer, 1955; Federer, 1957; Coons, 1957).

Wilkinson (1960) made a comparison of missing plot value procedures based on the least squares and covariance methods. Smith (1957) and Hoyle (1971) presented useful references dealing with use of analysis of covariance in problems related to missing observations.

3. MATERIALS AND METHODS

3.1 Materials

Synthesized data are often used for studying the properties of new statistical techniques. However, for developing methods of handling the effects of missing plants, actual data must be used because they reflect the environmental variations to which agricultural and forestry experiments are subjected. The author was able to obtain a number of data sets from various sources but many were not suitable for this study. The following properties were required of data sets to be considered for further study: (1) plants should be reasonably close together so that competition exists between adjacent plants; (2) the experimental design should be relatively simple and straightforward; (3) measurements should have been taken by individual plant; (4) the missing plants should be distributed rather evenly over the experimental area; and (5) there should be a variety of configurations of missing plants (e.g., singles, pairs, triplets, etc.). It was also necessary to have background information about how the experiment was carried out and a record of unusual events which might have affected the site.

The following data sets met most of the above requirements: DATA LOBLOLLY PINE 1, DATA LOBLOLLY PINE 2, DATA OIL PALM, and DATA PEACH. These data came from experiments which were designed for other purposes but the occurrence of missing plants rendered them useful for these missing plant studies. Descriptions of the data sets are as follows:

DATA LOBLOLLY PINE 1. This was set No. 7 from a heritability study on loblolly pine (*Pinus taeda* L.) conducted under the sponsorship of the International Paper Company and the North Carolina State University-Industry

Cooperative Tree Improvement Program. For details about these data, see Stonecypher et al. (1973).

Location: Decatur Co., Georgia.

Established: 1961.

Data recorded: 1976.

Source of data: Dr. Robert Weir, N. C. State University-Industry Cooperative Tree Improvement Program, Raleigh, N. C.

Characteristics of this experiment were:

Design: Randomized complete block.

Treatments: Twenty-eight open-pollinated families of loblolly pine identified as follows:

4E	28C	48B	39
17A	29A	48E	46
21A	32C	50C	48
21B	33C	54E	49
21E	38E	62B	59
24A	39A	12	64
24C	44B	15	65

Replications: Three (after eliminating a fourth from the original experiment which was not available for analysis).

Experimental unit: A 5 x 5 tree square plot. Spacing between plots and between trees within plots was eight feet (2.44 m). There were no border rows between plots in the same replication.

Measurements:

1. HT = total height of the tree to the nearest foot (30.5 cm).
2. DBH = diameter at breast height [4.5 feet (1.37 m) from the ground] measured on the outside bark to the nearest 0.1 inch (2.54 dm).

3. VOL = volume outside bark in cubic feet (1 ft^3 = 0.028 m^3), computed as $\hat{VOL}$ = 0.31995 + 0.00249 (DBH x DBH x HT).

DATA LOBLOLLY PINE 2. These data were recorded from a progeny test experiment on loblolly pine sponsored by the Weyerhaeuser Company and the North Carolina State University-Industry Cooperative Tree Improvement Program.

Location: Martin Co., North Carolina.

Established: 1964.

Data recorded: 1974.

Source of data: Dr. Robert Weir, N. C. State University-Industry Cooperative Tree Improvement Program, Raleigh, N. C.

Characteristics of this experiment were:

Design: Randomized complete block.

Treatments: Forty open-pollinated families of loblolly pine identified as follows:

90P	9CC	1200P	120CC
160P	16CC	1210P	121CC
350P	35CC	1220P	122CC
440P	44CC	1260P	126CC
580P	58CC	1270P	127CC
800P	80CC	1280P	128CC
1030P	103CC	1290P	129CC
1060P	106CC	1300P	130CC
1170P	117CC	1400P	140CC
1180P	118CC	1440P	144CC

Replications: Four.

Experimental unit: A one-row plot 25 trees long. Spacing between plots and trees within plots was nine feet (2.74 m). There were no border rows between plots in the same replication.

Measurements:

1. HT = total height of the tree to the nearest 0.1 foot (3.05 cm).

2. DBH = diameter at breast height (4.5 feet from the ground) measured on the outside bark to the nearest 0.1 inch.

3. VOL = volume outside bark in cubic feet computed as $\hat{VOL}$ = 0.03371 + 0.00196128 (DBH x DBH x HT).

DATA OIL PALM. These data were brought to my attention by Dr. L. A. Nelson, my advisor, who was acting as a consulting statistician for the analysis of the data from this experiment. The experiment was conducted at the Santo Domingo Experiment Station of INIAP (National Institute of Agricultural Investigations), Santo Domingo, Ecuador.

Location: Field 13D, experiment SD-Pal-02.011.68, Santo Domingo Experiment Station at INIAP, Santo Domingo, Ecuador.

Established: 1968.

Data recorded: 1972, 1973, 1974, and 1975.

Source of data: Dr. L. A. Nelson, Department of Statistics, N. C. State University, Raleigh, N. C.

Characteristics of this experiment were:

Design: Randomized complete block.

Treatments: Eight crosses of oil palm (<u>Elaesis guineensis</u> Jacq.) identified as follows:

212 x 430	430 x 250
137 x 430	254 x 250
212 x 268	137 x 250
113 x 268	137 x 113

Replications: Three.

Experimental unit: A one-row plot 24 trees long. Spacing between plots and trees within plots was 10 m. There were no border rows between plots in the same replication.

Measurements:

1. B_i = plant yield measured as the number of bunches per year i; i = 1972, 1973, 1974, and 1975.

2. Y_i = total weight of plant yield (kg) in year i.

3. MYB_i = mean weight per bunch for each plant in year i.

DATA PEACH. The experiment on peach (Prunus persica L.) was conducted by F. C. Correll, C. N. Clayton, and G. A. Cummings of the North Carolina Agricultural Experiment Station.

Location: Field F2D, Sandhills Research Station, Jackson Springs, North Carolina; same location for all years.

Established: 1966.

Data recorded: 1969, 1970, 1971, and 1972.

Source of data: Dr. G. A. Cummings, Department of Soil Science, N. C. State University, Raleigh, N. C.

Characteristics of this experiment were:

Design: A split-split-split-split-split plot design with the third and fourth split factors stripped over the other plots.

Treatments: The treatments applied to the various plots were as follows:

1st split: varieties (Candor, Redskin), randomized within replication.

2nd split: 2^2 factorial, nitrogen (normal, 1/2 normal), x post-plant soil fumigation (no Nemagon, Nemagon applied in 1968 and every two years thereafter) randomized within varieties.

3rd split: spacing between trees within rows, 12 and 20 feet (3.66 and 6.10 m, respectively), randomized and stripped over varieties.

4th split: pre-plant soil fumigation (D-D, no D-D), randomized within spacing and stripped over variety.

5th split: 2^2 factorial, rootstock (Lovell, Nemaguard) x pruning time (March, November), randomized within replication x variety x nitrogen x post-plant fumigation x spacing x pre-plant fumigation.

Replications: Three, spaced 20 feet apart.

Experimental unit: A one-row plot of two trees. Spacing between rows was 20 feet. There was a border row around the experimental field and between spacing levels.

Measurements:

1. TRAR = growth of the tree as measured by annual increases expressed in square inches (1 in.^{2} = 6.45 cm^2) of basal cross-sectional area of tree trunk approximately one foot from the soil line.
2. TOH = yield in pounds (1 lb = 0.45 kg) of fruit at harvest.
3. NOFR = number of fruit harvested per tree. This variable was subject to gross errors and thus its results were considered unreliable and they were not used in this study.[1]

3.2 Methods

The methods section is divided into two parts. The first one deals with the derivation and designation of the covariates and the other with the statistical analysis which was used.

3.2.1 Derivation and Designation of the Covariates

Consideration of what is meant by a missing plant and a standing plant is necessary for subsequent discussions. A missing plant is one

[1]Personal communication, Dr. C. N. Clayton, Department of Plant Pathology, N. C. State University at Raleigh.

that is not present (an empty space is in its place). Dead plants are not missing plants because they may produce competitive effects in the form of shade. In time, a dead plant may be absent from the field before the end of the experimental period and thus become a missing plant. A standing plant, whether or not it provides reliable data, cannot be considered a missing plant (e.g., a standing plant may be favored by plant-site variation and provide unreliable observations which have to be omitted from analysis because of gross errors). Injured and dwarf plants, although standing, may affect surrounding standing plants in the same way as do missing plants. Therefore, their responses should be omitted from statistical analysis in the same way as are those of missing plants. Hereafter, it should be understood that missing plants are those not present plus those standing plants taken as missing plants.

Missing plants and standing plants occur interspersed throughout a field. Each standing plant may be surrounded by missing plants from the same plot or different plots than the plot that contains the standing plant; also, missing plants in guard rows (if present) affect the responses of plants in the test area even if all plants in the test area are present and, therefore, they should be taken into account when developing missing plant procedures.

Because of some environmental factors like fertility gradients, direction and duration of solar radiation, and air circulation in the experimental field, the location of missing plants in relation to a standing plant seems to be an important factor for measuring the effects of missing plants. Hindi (1962) found that the effects of missing hills were more noticeable on hills located to the north or west than on hills located to the south or east, respectively, of the missing hills. Also, missing plants may affect a standing plant differently as the number of missing plants around the standing plant increases; this is reasonable because the wider the gap the more growing space and other growth factors will be available to the plants adjacent to the gap and, of course, less competition is expected as plants are more separated from each other. Therefore, a standing plant is expected to alter its response for each additional missing plant that occurs in its surrounding. For these reasons, it seems important to know how to quantify these effects of number of missing plants surrounding a standing plant so that proper adjustment be made on the response variable to improve the accuracy of treatment comparisons.

This thesis is written assuming a square planting pattern. The location of plants (missing or not) surrounding a standing plant will be designated as follows:

Vertical locations = v_1, v_2.

Horizontal locations = h_1, h_2.

Diagonal locations = d_1, d_2, d_3, d_4.

The locations for square planting are shown in Figure 3.1.

The number of missing plants along any location may be none, one, or more than one. Nevertheless, because diagonal distances are greater than vertical or horizontal distances, the effects of more than one missing plant along a diagonal are not considered to affect a standing plant. Certain locations are omitted from the analysis if the standing plant used to derive the covariates falls in a border row.

Two major covariate definition approaches were used: Approach (a) covers the possibility of significance of the missing plant effects upon

d_1	v_1	d_2
h_1	p	h_2
d_4	v_2	d_3

v_1 = a vertical location; $i = 1, 2$.

h_j = a horizontal location; $j = 1, 2$.

d_k = a diagonal location; $k = 1, 2, 3, 4$.

p = a standing plant for which the covariates are measured.

Figure 3.1 Locations of missing plants surrounding a standing plant for a square planting pattern location. A location must have the same, or practically the same, orientation for all replications in relation to some environmental factors like solar radiation and air circulation in the field. The experimenter needs to consider just one orientation [e.g., from north (v_1) to south (v_2) for vertical locations] because the others will be determined in relation to it. Approach (b) does not account for variation due to location. Instead, it considers the same effects regardless of orientation for all of the missing plants used to define a particular covariate. Some covariates are indicators of the presence (value = 1) or absence (value = 0) of missing plants; the others define the number of and closeness or location of the surrounding missing plants to a standing plant. The latter covariates are indeed more informative and are expected to show better results.

The approaches are:

1. Dependent upon the location of missing plants in relation to a standing plant.

   1.1 The covariates for a standing plant with respect to the vertical locations (v_1, v_2) are:

   V = 1 if missing plants exist at one or both vertical locations; 0 if missing plants do not exist at either vertical location.

   V1 = 1 if missing plants exist at one vertical location only; 0 if missing plants exist at both vertical locations.

   V2 = 1 if missing plants exist at both vertical locations; 0 if missing plants exist at one vertical location only.

VM = total number of missing plants at one or both vertical locations.

VM1 = total number of missing plants at one vertical location only.

VM2 = total number of missing plants at both vertical locations.

1.2 The covariates for a standing plant with respect to the horizontal locations (h_1, h_2) are: H, H1, H2, HM, HM1, and HM2; for the definition of these covariates, substitute H for V and horizontal for vertical in the above definitions.

1.3 The covariates for a standing plant with respect to the diagonal locations (d_1, d_2, d_3, 3_4) are:

D = 1 if missing plants exist at at least one diagonal location; 0 if missing plants do not exist at any of the four diagonal locations.

DM = total number of missing plants at all four diagonal locations.

2. Independent of the location of missing plants in relation to a standing plant. The covariates for a standing plant, p, are:

C = 1 if missing plants exist surrounding p; 0 if missing plants do not exist surrounding p.

TvM = total number of missing plants at a distance v missing plants away from p for $v \leq (\ell - 1)$.

= total number of missing plants at a distance $\geq$ v missing plants away from p for $v = \ell$, where ℓ is the distance measured as the number of missing plants away from p at which the effect of more than ℓ missing plants away from p is considered the same as that of ℓ missing plants away from p. In this thesis, $\ell = 4$ for all sets of data used in this study. For any particular data set, a researcher will need to depend upon his experience with the crop involved in determining ℓ.

TM = total number of missing plants surrounding p. TM may be computed by direct count, or as

= VM + HM + DM, or as

$$= \sum_{v=1}^{\ell} TvM.$$

SvM = total number of sets of v missing plants surrounding p for $v \leq (\ell - 1)$. For v = 1, each diagonal missing plant is considered as a set of one missing plant.

= total number of sets of $\geq$ v missing plants surrounding p for $v = \ell$.

SvM can also be expressed as a function of the number of missing plants by TSvM = v(SvM).

PT1M = total number of missing plants at a distance one missing plant away from p, in sets of more than one missing plant; it may be computed as

= T1M - S1M.

T1MT2M = T1M + TM2.

T3MT4M = T3M + T4M.

S1MS2M = S1M + S2M.

S3MS4M = S3M + S4M.

$LTM = \ell n(TM + 1)$, where ℓn refers to the natural logarithm.

$LT1M = \ell n(T1M + 1)$.

To study the effects of missing plants on the diagonal locations, DM is subtracted from each of S1M and T1M. In this way the effects due to vertical and horizontal missing plants are separated from those due to the diagonal missing plants. Therefore,

NS1M = S1M - DM.

NT1M = T1M - DM.

To ease calculation of TvM and SvM, several auxiliary co-variates were constructed for vertical, horizontal, and diagonal locations (e.g., VT1M = T1M on vertical locations; i.e., total number of missing plants on the vertical locations at a distance one missing plant away from p). Therefore,

T1M = VT1M + HT1M + DM.

TvM = VTvM + HTvM; v = 2, 3, 4.

S1M = VS1M + HS1M + DM.

SvM = VSvM + HSvM; v = 2, 3, 4.

Using this type of auxiliary covariate,

MMT = VMMT + HMMT where

VMMT = total number of missing plants between standing plants excluding the end missing plants, from sets of more than three missing plants on the vertical locations.

= 0 if sets of 0, 1, or 2 missing plants on the vertical locations surround p.

HMMT is defined in a similar manner to that of VMMT for horizontal locations.

All of these covariates take on the value = 0 when no missing plants surround a standing plant; moreover, missing plants are counted as missing observations for all response variables and covariates.

To illustrate the covariate designations with some examples, let us consider Figure 3.2 where some selected standing plants, p_i, are presented with and without surrounding missing plants; the covariates for these standing plants take on the values given in Table 3.1.

3.2.2 Statistical Analysis

The sets of covariates included in the model for adjustment are presented below. In each set, the covariates are ordered from left to right by their expected importance in contributing to the bias. Obviously, no two sets may be in the same model. The covariates in each set are:

Set no.	Covariates
1	C.
2	V, H, HV, D, DV, DH, DHV.
3	V1, V2, H1, H2, H1V1, H1V2, H2V1, H2V2, D, DV1, DV2, DH1, DH2, DH1V1, DH1V2, DH2V1, DH2V2.
4	VM, HM, HMVM, DM, DMVM, DMHM, DMHMVM.
5	VM1, VM2, HM1, HM2, HM1VM1, HM1VM2, HM2VM1, HM2VM2, DM, DMVM1, DMVM2, DMHM1, DMHM2, DMHM1VM1, DMHM1VM2, DMHM2VM1, DMHM2VM2.
6	TM.
7	TM, TM*TM.

x	x	x
x	p_1	x
x	x	x

x	x	x	x	x	x	x
x	x	o	x	x	x	x
x	x	p_4	x	o	o	x
x	o	o	x	x	x	x
x	x	o	x	x	x	x
x	x	o	x	o	x	x
x	x	o	x	x	x	x
x	x	x	x	x	x	x

x	x	x	x	x
x	p_2	o	o	x
x	o	o	x	x
x	x	x	x	x

x	x	x	x	x
x	x	o	x	x
x	x	o	x	x
x	x	o	o	x
x	o	p_3	o	x
x	x	x	x	x

x	x	x	x	x	x	x	x
x	x	x	o	x	o	o	x
x	o	o	o	o	o	p_5	x
x	x	x	x	x	o	o	x
x	x	x	o	x	x	o	x
x	x	x	x	x	x	x	x

x = a standing plant.

o = a missing plant.

p_i = a standing plant for which covariates are measured; i = 1,2,3,4,5.

Figure 3.2 Examples of some possible distributions of missing plants surrounding a standing plant

Table 3.1 Values taken by the covariates measured on the standing plants (p_i, i = 1, 2, 3, 4, 5) in Figure 3.2

Covariate	Standing plant				
	p_1	p_2	p_3	p_4	p_5
C	0	1	1	1	1
V	0	1	1	1	1
V1	0	1	1	0	0
V2	0	0	0	1	1
VM	0	1	3	5	3
VM1	0	1	3	0	0
VM2	0	0	0	5	3
H	0	1	1	0	1
H1	0	1	0	0	1
H2	0	0	1	0	0
HM	0	2	2	0	5
HM1	0	2	0	0	5
HM2	0	0	2	0	0
D	0	1	1	1	1
DM	0	1	1	1	2
TM	0	4	6	6	10
T1M	0	3	4	3	5
T2M	0	1	1	1	2
T3M	0	0	1	1	1
T4M	0	0	0	1	2
T1MT2M	0	4	5	4	7
T3MT4M	0	0	1	2	3
S1M	0	2	3	2	3
S2M	0	1	0	0	1
S3M	0	0	1	0	0
S4M	0	0	0	1	1
S1MS2M	0	3	3	2	4
S3MS4M	0	0	1	1	1
PT1M	0	1	1	1	2
MMT	0	0	1	2	3
NT1M	0	2	3	2	3
NS1M	0	1	2	1	1

Set no.	Covariates
8	T1M, T2M, T3M, T4M.
9	S1M, S2M, S3M, S4M.
10	S1M, PT1M, T2M, T3M, T4M.
11	T1MT2M, T3M, T4M.
12	T1MT2M, T3MT4M.
13	S1MS2M, S3M, S4M.
14	S1MS2M, S3MS4M.
15	T1M, MMT.
16	NT1M, DM.
17	NS1M, PT1M, DM.
18	NS1M, DM.
19	LTM.
20	LT1M.

It should be pointed out that in sets 2, 3, 4, and 5, there exist terms for the interaction among covariates. For example,

$$HV = H*V;\ H1V1 = H1*V1;\ HMVM = HM*VM;\ \text{and}\ HM1VM1 = HM1*VM1.$$

Sets 2-5 include the covariates depending upon the location of missing plants whereas sets 1 and 6-20 do not depend upon location. The covariate C (set 1) is a general covariate expected to express whether the effects of missing plants exist at all, regardless of where the missing plants surrounding a standing plant are. Subsequently, V (set 2) is used to determine whether missing plants at the vertical locations have some effects on adjacent standing plants. Since V = V1 + V2, V1 and V2 (set 3) express the relative importance of missing plants on one or both vertical locations, respectively. Similar arguments are given for H, H1, H2, and D. If the number of missing plants is taken into consideration, TM (set 6), VM (set 4), and VM1, VM2 (set 5) are used in a similar manner as C, V, and V1, V2, respectively. The same is true for HM, HM1, HM2, and DM. The quadratic effect of TM is studied in set 7. The inclusion of T1M, T2M, T3M, and T4M (set 8) allows the study of the importance of missing plants according to their closeness to a standing plant in the order T1M-T4M where T1M represents those missing plants closest to the standing plant. The inclusion of S1M, S2M, S3M, and S4M (set 9) is useful to test the effects of groups of 1, 2, 3, and 4 or more missing plants, respectively. T1MT2M (sets 11 and 12) measures the effects due to each additional missing plant at a distance one or two missing plants away from a standing plant, i.e., the effects of T1M are considered equal to those of T2M. T3MT4M (set 12) for three and four or more missing plants is defined in a similar manner to that of T1MT2M. S1MS2M (sets 13 and 14) measures the effects due to each additional set of one or two missing plants surrounding a standing plant, i.e., it assumes that the effects of S1M equal the effects of S2M. S3MS4M (set 14) for sets of three and four or more missing plants is defined in a similar manner to that of S1MS2M. In set 15, MMT measures the effects of missing plants (excluding T1M) which are between standing plants; T1M is investigated in more detail in set 10 (T1M = S1M + PT1M), and in sets 16 (T1M = NT1M + DM) and 17 (T1M = NS1M + PT1M + DM) especially to study the effects of DM. DM is also studied in set 18 where S1M = NS1M + DM. In set 19, LTM = ℓn (TM + 1) is the natural logarithm of (TM + 1) where TM is increased by one because the value TM = 0 is not in the domain of definition of the logarithmic function. LT1M = ℓn(T1M + 1) follows the same reasoning as LTM.

It should be kept in mind that we are dealing with a missing observations problem which does not have the straightforward analysis of data from a balanced design. Therefore, a regression procedure (Service, 1972) was used for an exact analysis of the data. (This would be referred to as an analysis of covariance approach.) The relative importance of the individual covariates was evaluated by finding the individual extra sum of squares in the regression procedure. Hence, the order of introduction of the covariates into the model was relevant because each additional covariate introduced into the model was adjusted for all class variables and covariates already in the model.

To illustrate the extra sum of squares (or sequential sum of squares) principle by means of an example involving two covariates, NT1M and DM, in a randomized complete block design, the regression sum of squares will be partitioned into the following components:

1. SS(REP): The sum of squares explained using replications (REP) as the only set of predictor variables in the model.
2. SS(TRT/REP): The extra sum of squares in the response variable due to treatments (TRT) in addition to REP.
3. SS(REP*TRT/REP, TRT): The extra sum of squares in the response variable due to the interaction REP*TRT in addition to REP and TRT.
4. SS(NT1M/REP, TRT, REP*TRT): The extra sum of squares in the response variable due to the covariate NT1M in addition to REP, TRT, and REP*TRT.
5. SS(DM/REP, TRT, REP*TRT, NT1M): The extra sum of squares in the response variable due to the covariate DM in addition to REP, TRT, REP*TRT, and NT1M.

For another example, T1M, T2M, T3M, and T4M were introduced in that order after REP*TRT in the above experimental design because this was the order of importance which the covariates were expected to have in evaluating the effects of missing plants.

To test whether the addition of any specific independent covariate significantly contributes to the bias per missing plant on the response variable, the F-test was used. Results of the test may be used to determine which covariates are not important in estimating the bias thus enabling one to reduce the set of possible independent covariates to the "best" set of estimators of bias.

The statistical significance of a covariate will be an indication of its importance in detecting the effects of missing plants whereas the value of its regression coefficient will be an estimation of the average amount of compensation (bias) for each missing plant. This bias per missing plant will be positive or negative if the regression coefficient is positive or negative, respectively.

To adjust treatment and experimental error sums of squares of variation to a constant level of NT1M, say, this covariate was introduced into the model immediately after REP*TRT.

To test the significance of treatment effects after adjusting for NT1M, the following F-test was computed:

$$F = \frac{\text{MST after adjustment for REP and NT1M}}{\text{MSE after adjustment for REP, NT1M, and TRT}}$$

where MST is the treatment mean square and MSE is the error mean square.

One would like to find a single model with as few covariates as possible which would explain the bias due to missing plants and which is easy to fit to experimental data. Therefore, the following steps were taken in this study: (1) all of the covariance models suggested in this thesis were compared among themselves and with the model without covariates, (2) the model which fit the data best and yet was easy to fit was chosen for subsequent analysis of covariance, (3) the covariates which were most important in measuring the missing plant effects were selected from the model obtained in (2), and (4) the treatment effects were adjusted for the best covariates.

An illustration of the elimination of bias through adjustment in connection with a covariance analysis follows. For this discussion, the following definitions are needed:

$NT1M_{ijk}$ = total number of missing plants nearest to the $(i,j,k)^{th}$ standing plant, excluding diagonal missing plants.

$NT1M_{.j.}$ = total number of missing plants nearest to the standing plants for the j^{th} treatment, excluding diagonal missing plants.

$NT1M_{...}$ = total number of missing plants nearest to the standing plants in the entire experiment, excluding diagonal missing plants.

$n_{.j.}$ = total number of standing plants for the j^{th} treatment.

$n_{...}$ = total number of standing plants in the entire experiment.

$\overline{NT1M}_{.j.}$ = $NT1M_{.j.}/n_{.j.}$, mean number of missing plants nearest to a standing plant for the j^{th} treatment, excluding diagonal missing plants.

$\overline{NT1M}_{...}$ = $NT1M_{...}/n_{...}$, mean number of missing plants nearest to a standing plant in the entire experiment, excluding diagonal missing plants.

b = bias per missing plant nearest to a standing plant as measured by the regression coefficient of the response variable regressed on NT1M.

y_{ijk} = the $(i,j,k)^{th}$ tree response.

$y_{\cdot j \cdot}$ = the j^{th} treatment total.

$\bar{y}_{\cdot j \cdot}$ = the j^{th} treatment mean.

$y_{\cdots}$ = the overall experiment total.

$\bar{y}_{\cdots}$ = the overall experiment mean.

Therefore, from these definitions given above one could compute:

1. Total bias per tree = $bNT1M_{ijk}$
2. Total bias per treatment = $bNT1M_{\cdot j \cdot}$
3. Average bias per treatment = $b\overline{NT1M}_{\cdot j \cdot}$
4. Total bias for the entire experiment = $bNT1M_{\cdots}$
5. Average bias for the entire experiment = $b\overline{NT1M}_{\cdots}$

   To adjust for the bias (afb) to what would had been the response if there were no missing plants it will be assumed that a tree response was altered by the missing plants surrounding it, i.e.,

$$y_{ijk} = y_{ijk} \text{ (if no missing surrounding plants) } + bNT1M_{ijk}$$

   Therefore, to adjust for the bias (afb) we proceed as follows:

6. To adjust a tree response:

$$y_{ijk}(afb) = y_{ijk} - bNT1M_{ijk}$$

7. To adjust a treatment total

$$y_{\cdot j \cdot}(afb) = y_{\cdot j \cdot} - bNT1M_{\cdot j \cdot}$$

8. To adjust a treatment mean:

$$\bar{y}_{\cdot j \cdot}(afb) = \bar{y}_{\cdot j \cdot} - b\overline{NT1M}_{\cdot j \cdot}$$

9. To adjust the overall experiment total:

$$y_{\cdot\cdot\cdot}(afb) = y_{\cdot\cdot\cdot} - bNT1M_{\cdot\cdot\cdot}$$

10. To adjust the overall experiment mean:

$$\bar{y}_{\cdot\cdot\cdot}(afb) = \bar{y}_{\cdot\cdot\cdot} - b\overline{NT1M}_{\cdot\cdot\cdot}$$

11. To adjust treatment means to an equal average number of missing nearest surrounding plants:

$$\bar{y}_{\cdot j \cdot}(adj) = \bar{y}_{\cdot j \cdot} - b(\overline{NT1M}_{\cdot j \cdot} - \overline{NT1M}_{\cdot\cdot\cdot})$$

12. The expected treatment total if no bias were present in the entire experiment:

$\bar{y}_{\cdot j \cdot}(afb)$ x total number of standing plants for the j^{th} treatment if there were no missing plants.

13. The expected experiment total if no bias were present in the entire experiment:

$\bar{y}_{\cdot\cdot\cdot}(afb)$ x total number of standing plants for the entire experiment if there were no missing plants.

14. The average bias for the entire experiment expressed as a percentage of the adjusted overall experiment mean:

$$\frac{\bar{y}\ldots - \bar{y}\ldots(\text{afb})}{\bar{y}\ldots(\text{afb})} \times 100$$

This quantity will reflect a percent increase or decrease with respect to the adjusted overall experiment mean depending upon whether the bias is positive or negative, respectively.

The above procedure would apply in a similar manner to other designs and sets of covariates.

Two equivalent measures of the effectiveness of the covariance analysis approach are given by the percent reduction in the mean square for experimental error (reduction in MSE) due to the covariate(s) which is computed as

$$\text{Reduction in MSE} = \left(1 - \frac{\text{MSE after adjustment}}{\text{MSE before adjustment}}\right) \times 100,$$

and the relative efficiency on a percent basis computed as

$$\text{Relative efficiency} = \frac{\text{MSE before adjustment}}{\text{MSE after adjustment}} \times 100.$$

These two measures are positively associated because the increase in reduction in MSE implies an increase in relative efficiency. Also, the coefficients of variation (C.V.) on a percent basis were computed before and after adjusting the MSE for the missing plant effects variation. The C.V.'s of all models were then compared.

The prime interest in this study was to estimate the missing plant effects on adjacent standing plants and to adjust the response variables by those effects rather than to find estimates of the values for the missing plant. After the response variables have been adjusted for the effects of missing plants measured by the regression coefficient(s) of the covariate(s) defined in this thesis, the data should be analyzed to account for stand differences. The covariance analysis of the plot responses adjusted for the missing plant effects using standing plants as the covariate may be carried out. The analysis of plot means adjusted for the missing plant effects could also be conducted.

For the PEACH DATA, it is necessary to explain the meaning of some terms in the analysis of variance.

NP = 2^2 factorial of nitrogen*post-plant fumigation with treatment combinations:

11 = (1/2 normal nitrogen; no Nemagon) ≡ 1/2 normal nitrogen.

12 = 1/2 normal nitrogen; Nemagon.

21 = (normal nitrogen; no Nemagon) ≡ normal nitrogen.

22 = normal nitrogen; Nemagon.

RP = 2^2 factorial of rootstock*pruning time with treatment combinations:

11 = Lovell; November.

12 = Lovell; March.

21 = Nemaguard; November.

22 = Nemaguard; March.

S = spacing with levels:

1 = 12 feet.

2 = 20 feet.

PRE = pre-plant fumigation with levels:

1 = no D-D.

2 = D-D.

The mean squares associated with the sources of variation called error terms E( )'s, were computed by pooling the corresponding sum of squares of the terms shown below to the right of the colon sign and then dividing by the pooled degrees of freedom:

E(a) : REP*NP; REP = replication or block effects.

E(b) : REP*S.

E(c) : REP*S*NP.

E(d) : REP*PRE; REP*S*PRE.

E(e) : REP*NP*PRE; REP*NP*S*PRE.

E(f) : REP*RP; REP*NP*RP; REP*S*RP; REP*NP*S*RP; REP*PRE*RP; REP*NP*PRE*RP; REP*S*PRE*RP; REP*NP*S*PRE*RP.

4. RESULTS AND DISCUSSION

The mean squares attributable to the covariates in a particular set are presented in Tables 4.1, 4.2, 4.12, 4.13, 4.24, and 4.34. In these tables the probability of a larger F-value than the one computed (Prob > F) is presented for those covariates which were significant at the P = 0.01 or P = 0.05 levels, and for those covariates with (Prob > F) < 0.10.

DATA LOBLOLLY PINE 1. The statistical significance of the covariates is presented in Tables 4.1 and 4.2. In Table 4.2, C was not significant for the variables HT, DBH, and VOL which implies that missing plants had no detectable effects on these variables. This was not the case, as some covariates dependent upon location (Table 4.1) and others independent of location (Table 4.2) were significant.

For HT, the effects of missing plants were significant when it was measured only by some of the covariates dependent upon location. The covariate HV was significant at the P = 0.01 level and its interaction component H2V2 was significant at the P = 0.05 level. The covariate HMVM and its interaction component HM2VM2 were significant at the P = 0.05 level. The Prob > F value for H2V2 was 0.0148 and that for HM2VM2 was 0.0151; this does not seem to indicate a large difference in predictability between them. Hence, HT was more affected by the combined effects of missing plants which occurred in both vertical and both horizontal locations.

For DBH and VOL (Table 4.1) the covariates HV, H1V1, and D were significant at the P = 0.05 level. For VOL, DH1 was significant at the P = 0.05 level. The covariate HMVM was significant at the P = 0.01

Table 4.1 Mean squares attributable to covariates dependent upon location of missing plants in relation to a standing plant and results of tests of significance of same for DATA LOBLOLLY PINE 1

Covariate set	Mean square			Prob > F†		
	HT	DBH	VOL	HT	DBH	VOL
Set 2						
V	156.17	8.23	31.97	0.0409	0.0590	0.0257
H	86.58	29.45	99.70		0.0004	0.0001
HV	248.66	14.71	36.15	0.0099	0.0116	0.0177
D	113.52	13.92	34.17		0.0141	0.0211
DV	0.49	0.64	3.18			
DH	46.10	7.19	31.85			0.0259
DHV	37.14	0.24	2.61			
Set 3						
V1	143.37	5.34	30.52	0.0502		0.0293
V2	13.80	3.79	1.99			
H1	112.46	24.49	78.65		0.0012	0.0005
H2	1.30	5.63	20.84			
H1V1	130.82	11.81	29.90		0.0239	0.0310
H1V2	2.67	1.30	1.89			
H2V1	99.72	6.48	9.01			
H2V2	221.97	2.84	8.50	0.0148		
D	103.11	13.36	33.25		0.0163	0.0230
DV1	5.50	0.02	0.01			
DV2	44.26	4.70	19.17			
DH1	32.67	7.04	38.37			0.0146
DH2	20.19	0.29	0.06			
DH1V1	17.28	1.04	2.10			
DH1V2	81.95	1.20	0.91			
DH2V1	5.65	0.39	12.02			
DH2V2	9.04	1.35	2.41			
Set 4						
VM	162.79	6.63	14.56	0.0372		
HM	12.08	22.63	70.84		0.0018	0.0010
HMVM	260.82	16.31	27.92	0.0048	0.0080	0.0380
DM	12.51	5.34	10.13			
DMVM	6.40	1.75	2.48			
DMHM	26.90	4.19	8.10			
DMHMVM	0.31	0.94	3.90			

Table 4.1 (Continued)

Covariate set	Mean square			Prob > F†		
	HT	DBH	VOL	HT	DBH	VOL
Set 5						
VM1	117.53	5.44	20.06			
VM2	46.62	1.41	0.18			
HM1	36.13	14.94	46.57		0.0112	0.0074
HM2	2.43	7.89	24.73			0.0508
HM1VM1	95.23	14.90	28.02		0.0114	0.0377
HM1VM2	0.01	1.83	6.23			
HM2VM1	36.69	5.15	3.92			
HM2VM2	221.81	3.12	7.57	0.0151		
DM	10.75	5.39	10.60			
DMVM1	116.20	0.02	0.57			
DMVM2	90.27	5.84	18.32			
DMHM1	13.01	4.40	19.74			
DMHM2	9.35	0.15	1.21			
DMHM1VM1	5.32	4.69	8.46			
DMHM1VM2	10.40	0.08	1.26			
DMHM2VM1	3.60	0.01	0.86			
DMHM2VM2	15.60	0.59	0.01			

†A blank implies that (Prob > F) > 0.10.

level for DBH, and at the P = 0.05 level for VOL. Its interaction component HM1VM1 was significant at the P = 0.05 level for both DBH and VOL. No diagonal missing plant effects measured by DM were significant. The Prob F values for H1V1 were 0.0239 for DBH, and 0.0310 for VOL; those for HM1VM1 were 0.0114 for DBH, and 0.0377 for VOL. This indicates that the combined effects of missing plants on one horizontal and on one vertical location were well explained by either H1V1 or HM1VM1. Moreover, the significant effect of DH1 as a predictor of VOL suggests that the interaction of missing plants on the diagonal locations with those on just one horizontal location was important. From the results in Table 4.2, one observes that the effects of missing plants were well explained by the covariates independent of location which were significant at the P = 0.01 level. Furthermore, the Prob > F values were smaller than those for the covariates dependent upon location. For DBH and VOL, the significant effect of TM reflected a significant T1M effect. Comparing those sets that split the effect of T1M (10, 16, and 17), one sees that NT1M was the covariate of greatest importance. S1M explained most of the variation due to TM when sets of missing plants were taken into account and the effect of S1M was via NS1M. It was also observed that the diagonal effects measured by DM were not significant. It should be noted that the mean square for T2M in set 10 was zero. Both PT1M and T2M existed simultaneously in sets of more than one missing plant and, therefore, the definition of one implied the definition of the other. For this reason, it was not possible to separate both effects from each other and so PT1M has absorbed the effect of T2M.

Table 4.2 Mean squares attributable to covariates independent of location of missing plants in relation to a standing plant and results of tests of significance of same for DATA LOBLOLLY PINE 1

Covariate set	Mean square			Prob > F†	
	HT††	DBH	VOL	DBH	VOL
Set 1					
C	11.08	6.12	17.05		
Set 6					
TM	4.26	32.15	81.05	0.0002	0.0004
Set 7					
TM	4.26	32.15	81.05	0.0002	0.0004
TM*TM	0.23	2.12	0.19		
Set 8					
T1M	3.96	34.11	95.08	0.0001	0.0001
T2M	120.30	0.32	0.02		
T3M	0.23	0.07	0.01		
T4M	38.67	3.05	4.29		
Set 9					
S1M	34.36	27.34	81.46	0.0006	0.0004
S2M	67.44	5.05	10.93		
S3M	41.06	0.22	0.33		
S4M	1.68	4.03	4.79		
Set 10					
S1M	34.36	27.34	81.46	0.0006	0.0004
PT1M	89.91	7.10	13.63		
T2M	0.00	0.00	0.00		
T3M	0.23	0.07	0.01		
T4M	38.67	3.05	4.29		
Set 11					
T1MT2M	2.82	32.59	85.88	0.0002	0.0003
T3M	21.91	0.08	1.47		
T4M	37.73	3.01	4.20		
Set 12					
T1MT2M	2.82	32.59	85.88	0.0002	0.0003
T3MT4M	4.32	0.15	0.14		
Set 13					
S1MS2M	7.86	32.24	92.38	0.0002	0.0002
S3M	39.18	0.22	0.33		
S4M	1.71	4.03	4.78		
Set 14					
S1MS2M	7.86	32.24	92.38	0.0002	0.0002
S3MS4M	21.44	2.11	2.70		
Set 15					
T1M	3.96	34.11	95.08	0.0001	0.0001
MMT	6.03	1.18	0.59		
Set 16					
NT1M	7.47	34.99	112.79	0.0001	0.0001
DM	25.75	6.58	13.42		

Table 4.2 (Continued)

Covariate set	Mean square			Prob > F†	
	HT††	DBH	VOL	DBH	VOL
Set 17					
NS1M	6.20	28.97	103.57	0.0004	0.0001
PT1M	103.32	6.20	12.53		
DM	19.74	6.47	12.36		
Set 18					
NS1M	6.20	28.97	103.57	0.0004	0.0001
DM	31.53	4.98	9.44		
Set 19					
LTM	5.57	24.33	67.00	0.0013	0.0003
Set 20					
LT1M	0.02	25.74	72.52	0.0009	0.0008

†A blank implies that (Prob > F) > 0.10.

††HT was not significant for any of the covariates.

In Table 4.3, all covariate sets reduced the MSE except set 4 for HT where a negative reduction was obtained, i.e., the MSE was slightly increased. The use of sets 2 and 3 gave larger reductions in MSE of about 3.3% for HT, 13% for DBH, and 10% for VOL. All other sets did not differ greatly among themselves and reduced the MSE about 1% for HT, 10.5% for DBH, and 9.5% for VOL.

The C.V.'s (Table 4.4) for the model without covariates were about 33%, 28%, and 62% for HT, DBH, and VOL, respectively, and they differed little from those for the models with covariates.

Because set 16, which included a significant NT1M term, differed slightly (about 2% reduction in MSE) from sets 2 and 3, a set with just NT1M was tested with the expectation that NT1M was sufficient to measure the missing plant effects.

In Table 4.5 are presented the mean squares and F-values before and after adjusting treatments for missing plants using NT1M. The reduction in mean square was greater for experimental error than for treatments and, consequently, the F-value for treatments was larger after the adjustment was performed than before adjustment. Treatment effects for DBH and VOL were significant after the adjustment by NT1M at the P = 0.05 level.

In Table 4.6 an increased precision is shown by the use of NT1M alone, and it was even larger than that given by the use of sets 2 and 3; the reductions in MSE were 14.11% for DBH and 11.26% for VOL.

The missing plant effects (bias) per missing plant surrounding a standing plant measured by the regression coefficient of NT1M as a predictor of each of HT, DBH, and VOL are shown in Table 4.7. The bias was negative (-0.09 feet) and not significant for HT at the P = 0.05

Table 4.3 Percent reduction in MSE and relative efficiency for adjustment of DATA LOBLOLLY PINE 1 using covariates in the 20 different sets

Set	Reduction in MSE (%)			Relative efficiency (%)		
	HT	DBH	VOL	HT	DBH	VOL
1	0.13	3.90	3.21	100.13	104.06	103.31
2	2.83	13.31	10.42	102.91	115.35	111.63
3	3.84	12.81	10.36	103.99	114.69	111.56
4	-0.24	10.04	8.12	99.76	111.16	108.84
5	2.35	9.93	7.59	102.41	111.02	108.22
6	0.32	11.22	8.85	100.32	112.64	109.71
7	0.74	10.53	8.82	100.74	111.78	109.67
8	0.99	10.81	8.76	101.00	112.12	109.60
9	0.07	10.68	8.69	100.07	111.96	109.52
10	0.99	10.81	8.76	101.00	112.12	109.60
11	0.39	11.03	9.07	100.40	112.40	109.97
12	0.70	11.38	9.11	100.70	112.84	110.02
13	0.77	10.38	8.53	100.78	111.56	109.32
14	0.95	10.75	8.60	100.96	112.05	109.41
15	1.00	10.82	8.66	101.01	112.14	109.48
16	1.03	10.02	8.14	101.04	111.14	108.86
17	0.36	10.08	7.75	100.36	111.21	108.40
18	1.12	6.38	4.96	101.13	106.81	105.22
19	0.49	10.46	8.46	100.49	111.68	109.24
20	0.83	9.69	8.01	100.83	110.73	108.71

Table 4.4 Coefficient of variation (C.V.) expressed in percent derived from MSE for models with and without covariates for DATA LOBLOLLY PINE 1

Model	C.V. (%)		
	HT	DBH	VOL
Without covariates	33.29	27.90	62.23
With covariates (sets)			
1	33.26	27.35	61.23
2	32.81	25.98	58.90
3	32.64	26.05	58.92
4	33.33	26.46	59.65
5	32.89	26.48	59.83
6	33.23	26.29	59.42
7	33.16	26.39	59.43
8	33.29	26.35	59.45
9	33.27	26.37	59.47
10	33.29	26.35	59.45
11	33.22	26.32	59.35
12	33.17	26.27	59.33
13	33.16	26.42	59.52
14	33.13	26.36	59.50
15	33.12	26.35	59.48
16	33.11	26.47	59.65
17	33.23	26.46	59.78
18	33.10	27.00	60.67
19	33.29	26.40	59.54
20	33.15	26.52	59.69

Table 4.5 Unadjusted and adjusted mean squares using NT1M as the covariate and results of tests of significance of same for DATA LOBLOLLY PINE 1

Variable	Source	d.f.	Mean square Unadjusted	Mean square Adjusted	F Unadjusted	F Adjusted
DBH	Replications	2				
	Treatments	27	5.56	5.30	1.71*	1.89*
	Exp. error	54	3.26	2.80		
	Sampling error	1613				
VOL	Replications	2				
	Treatments	27	17.87	17.57	1.64 n.s.	1.81*
	Exp. error	54	10.92	9.69		
	Sampling error	1613				

*Significant at P = 0.05; n.s. = nonsignificant at P = 0.05.

Table 4.6 Percent reduction in MSE and relative efficiency for adjustment of DATA LOBLOLLY PINE 1 using the covariate NT1M

Covariate	Variable	Reduction in MSE (%)	Relative efficiency (%)
NT1M	DBH	14.11	116.43
	VOL	11.26	112.69

Table 4.7 Simple regression coefficients of the response variables HT, DBH and VOL on the covariate NT1M and standard errors of same for DATA LOBLOLLY PINE 1

Covariate	Variable	Regression coefficient, b	Standard error of b	t for H_0: $\beta = 0$	Prob > |t|
NT1M	HT	-0.09346889	0.20962315	-0.44589	0.6557
	DBH	0.20225330	0.05209994	3.88203	0.0001
	VOL	0.36311174	0.08692373	4.17736	0.0001

level. Therefore, HT was not studied any further for these data. The bias for DBH was 0.20 inches and that for VOL was 0.36 cubic feet; their standard errors were about the same (0.05 inches, and 0.09 cubic feet, respectively).

The regression coefficients and the total and mean of NT1M per treatment and for the entire experiment (Table 4.8) were used in the computation of the results presented in Tables 4.9, 4.10, and 4.11.

Very small differences (less than 0.15 in magnitude for each of DBH and VOL) were observed between unadjusted and adjusted treatment means when the adjustment was to be an equal average number of missing plants surrounding a standing plant (Table 4.9).

The estimated bias for DBH (Table 4.10) was smaller than that for VOL (Table 4.11). The bias was positive and thus the response of each tree surrounded by missing trees was increased with respect to the response of a tree not having surrounding missing trees. The total bias over all plants affected by missing plants for treatments ranged from 5.0563 to 11.7307 inches for DBH and from 9.0778 to 21.0605 cubic feet for VOL.

For the entire experiment, the percentage of missing plants was 19.19%. NT1M had an overall effect on all of the standing plants surrounded by missing plants of 247.9625 inches and 445.1750 cubic feet for DBH and VOL, respectively. The average biases for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment were 2.3090% for DBH and 5.1933% for VOL.

DATA LOBLOLLY PINE 2. As shown in Table 4.12, the missing plant effects were successfully measured by some covariates at the P = 0.05 level. These effects were similar for HT and VOL with respect to the

Table 4.8 Averages of the covariate NT1M for each treatment in DATA LOBLOLLY PINE 1

Treatment	n†	NT1M	Mean of NT1M
4E	55	52	0.9455
17A	67	29	0.4328
21A	65	35	0.5385
21B	58	40	0.6897
21E	60	46	0.7667
24A	59	54	0.9153
24C	55	51	0.9273
28C	61	44	0.7213
29A	57	50	0.8772
32C	71	25	0.3521
33C	61	43	0.7049
38E	56	52	0.9286
39A	65	29	0.4462
44B	55	53	0.9636
48B	63	45	0.7143
48E	68	26	0.3824
50C	58	48	0.8276
54E	65	37	0.5692
62B	67	25	0.3731
12	58	52	0.8966
15	61	45	0.7377
39	58	58	1.0000
46	54	51	0.9444
48	54	48	0.8889
49	64	43	0.6719
59	59	55	0.9322
64	61	45	0.7377
65	62	45	0.7258
For the experiment	1697	1226	0.7225

†Number of standing plants. The percentage of missing plants for the entire experiment was 19.19%.

Table 4.9 Unadjusted and adjusted treatment means when the adjustment was to an equal average number of missing plants surrounding a standing plant using NT1M as the covariate for DATA LOBLOLLY PINE 1

Treatment	n†	DBH (inches)		VOL (cubic feet)	
		Unadjusted	Adjusted	Unadjusted	Adjusted
4E	55	6.68	6.63	6.12	6.04
17A	67	6.64	6.70	5.51	5.62
21A	65	6.81	6.85	5.88	5.95
21B	58	5.93	5.93	4.30	4.31
21E	60	6.97	6.96	6.22	6.20
24A	59	6.29	6.25	4.94	4.87
24C	55	5.83	5.79	4.30	4.23
28C	61	6.42	6.42	4.97	4.97
29A	57	6.55	6.52	5.61	5.55
32C	71	6.04	6.11	4.65	4.78
33C	61	6.58	6.59	5.95	5.96
38E	56	6.26	6.22	4.78	4.71
39A	65	6.03	6.09	4.55	4.65
44B	55	6.58	6.53	5.38	5.29
48B	63	6.57	6.57	5.37	5.37
48E	68	6.42	6.49	5.18	5.31
50C	58	6.35	6.33	4.97	4.93
54E	65	6.02	6.05	4.49	4.55
62B	67	6.57	6.65	5.69	5.81
12	58	6.56	6.52	5.39	5.33
15	61	6.86	6.86	5.95	5.95
39	58	6.43	6.38	5.19	5.09
46	54	6.65	6.60	5.59	5.51
48	54	6.16	6.13	4.72	4.65
49	64	6.55	6.56	5.59	5.61
59	59	6.85	6.81	5.89	5.81
64	61	7.03	7.03	6.15	6.15
65	62	6.63	6.63	5.43	5.43

†Number of standing plants.

Table 4.10. Estimated bias per treatment as measured by the covariate NT1M for the variable DBH in DATA LOBLOLLY PINE 1

Treat-ment	Bias† Total	Bias† Average	Mean Unadjusted	Mean Adjusted	Total if no bias present
	inches				
4E	10.5172	0.1912	6.6800	6.4888	486.6584
17A	5.8653	0.0875	6.6418	6.5543	491.5693
21A	7.0789	0.1089	6.8123	6.7034	502.7546
21B	8.0901	0.1395	5.9276	5.7881	434.1086
21E	9.3037	0.1551	6.9733	6.8182	511.3679
24A	10.9217	0.1851	6.2864	6.1013	457.5965
24C	10.3149	0.1875	5.8291	5.6416	423.1167
28C	8.8991	0.1328	6.4180	6.2852	471.3883
29A	10.1127	0.1774	6.5509	6.3735	478.0114
32C	5.0563	0.0712	6.0394	5.9682	447.6138
33C	8.6969	0.1426	6.5836	6.4410	483.0771
38E	10.5172	0.1878	6.2625	6.0747	455.6020
39A	5.8653	0.0902	6.0292	5.9390	445.4223
44B	10.7194	0.1949	6.5818	6.3869	479.0176
48B	9.1014	0.1445	6.5683	6.4238	481.7875
48E	5.2586	0.0773	6.4221	6.3448	475.8576
50C	9.7082	0.1674	6.3466	6.1792	463.4413
54E	7.4834	0.1151	6.0185	5.9034	442.7528
62B	5.0563	0.0755	6.5746	6.4991	487.4349
12	10.5172	0.1813	6.5552	6.3739	478.0402
15	9.1014	0.1492	6.8639	6.7147	503.6023
39	11.7307	0.2023	6.4328	6.2305	467.2910
46	10.3149	0.1910	6.6463	6.4553	484.1462
48	9.7082	0.1798	6.1611	5.9813	448.5989
49	8.6969	0.1359	6.5453	6.4094	480.7058
59	11.1239	0.1885	6.8542	6.6657	499.9244
64	9.1014	0.1492	7.0295	6.8803	516.0223
65	9.1014	0.1468	6.6258	6.4790	485.9252
For the experiment	247.9625	0.1461	6.4735	6.3274	13287.5018

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is 2.3090%.

Table 4.11 Estimated bias per treatment as measured by the covariate NT1M for the variable VOL in DATA LOBLOLLY PINE 1

Treat-ment	Bias† Total	Bias† Average	Mean Unadjusted	Mean Adjusted	Total if no bias present
	-----------------	-----------------	cubic feet	-----------------	-----------------
4E	18.8818	0.3434	6.1243	5.7810	433.5746
17A	10.5302	0.1572	5.5124	5.3552	401.6424
21A	12.7089	0.1955	5.8819	5.6864	426.4784
21B	14.5245	0.2504	4.2976	4.0472	303.5384
21E	16.7031	0.2784	6.2208	5.9424	445.6811
24A	19.6080	0.3323	4.9423	4.6100	345.7470
24C	18.5187	0.3367	4.2999	3.9632	297.2397
28C	15.9769	0.2619	4.9695	4.7076	353.0688
29A	18.1556	0.3185	5.6080	5.2895	396.7111
32C	9.0778	0.1279	4.6482	4.5203	339.0258
33C	15.6138	0.2560	5.9529	5.6969	427.2702
38E	18.8818	0.3372	4.7846	4.4474	333.5569
39A	10.5302	0.1620	4.5477	4.3857	328.9272
44B	19.2449	0.3499	5.3799	5.0300	377.2494
48B	16.3400	0.2594	5.3657	5.1063	382.9751
48E	9.4409	0.1388	5.1828	5.0440	378.2972
50C	17.4294	0.3005	4.9662	4.6657	349.9270
54E	13.4351	0.2067	4.4908	4.2841	321.3079
62B	9.0778	0.1355	5.6854	5.5499	416.2433
12	18.8818	0.3255	5.3938	5.0683	380.1189
15	16.3400	0.2679	5.9549	5.6870	426.5273
39	21.0605	0.3631	5.1861	4.8230	361.7241
46	18.5187	0.3429	5.5914	5.2485	393.6346
48	17.4294	0.3228	4.7152	4.3924	329.4326
49	15.6138	0.2440	5.5875	5.3435	400.7651
59	19.9711	0.3385	5.8876	5.5491	416.1830
64	16.3400	0.2679	6.1510	5.8831	441.2348
65	16.3400	0.2635	5.4316	5.1681	387.6038
For the experiment	445.1750	0.2623	5.3130	5.0507	10606.4058

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is 5.1933%.

covariates H and HM for which only H1 and HM1 were significant. For DBH, H and HM were marginally significant. Also, V was marginally significant for HT. However, the vertical, horizontal, and diagonal effects interacted as was shown by the significance of DHV for DBH and DMHMVM for HT and DBH. These two covariates were marginally significant for VOL. This interaction effect was measured successfully by the individual components DH1V2 for HT, and DH1V1 for DBH and VOL. There were no standing plants with surrounding missing plants on both vertical (v_1, v_2) and both horizontal (h_1, h_2) locations; therefore, H2V2 = 0 for every standing plant and mean squares for the terms H2V2, DH2V2, HM2VM2, and DMHM2VM2 could not be computed.

In Table 4.13 the covariates C and TM were significant for HT at the P = 0.01 level. Most of the effects of these covariates were via T1M. The effect of T1M was largely explained by NT1M in set 16 which was significant for HT and VOL at the P = 0.01 and P = 0.05 levels, respectively. NT1M was marginally significant for DBH with (Prob > F) = 0.0602. Although for DBH, LTM and LT1M were significant at the P = 0.05 level, their Prob > F values (0.0445 and 0.0500, respectively) differed little from that for NT1M. In set 10, the effect of T2M was absorbed by PT1M in the manner explained for DATA LOBLOLLY PINE 1.

In Tables 4.14 and 4.15, the sets of covariates are compared among themselves and with the model having no covariates. The reductions in MSE were about 3% for HT and about 2% for DBH and VOL. Sets 4 and 5 had the largest reductions in MSE (about 4%, 3%, and 2% for HT, DBH, and VOL, respectively), but they differed little from sets 16 and 19 (about 3%, 2%, and 2% for HT, DBH, and VOL, respectively). The C.V.'s (Table 4.15) for the model without covariates were about 46%, 51%, and 125% for

Table 4.12 Mean squares attributable to covariates dependent upon location of missing plants in relation to a standing plant and results of tests of significance of same for DATA LOBLOLLY PINE 2

Covariate set	Mean square			Prob > F†		
	HT	DBH	VOL	HT	DBH	VOL
Set 2						
V	41.87	0.71	0.20	0.0620		
H	75.22	3.63	1.09	0.0124	0.0635	0.0109
HV	10.55	0.44	0.19			
D	61.94	0.75	0.03	0.0232		
DV	13.52	0.30	0.01			
DH	19.90	4.94	0.68		0.0304	0.0434
DHV	38.47	5.59	0.63	0.0736	0.0213	0.0510
Set 3						
V1	38.19	0.79	0.23	0.0745		
V2	3.72	0.01	0.01			
H1	55.18	2.46	0.97	0.0320		0.0164
H2	24.45	1.58	0.12			
H1V1	1.50	0.01	0.04			
H1V2	12.98	0.90	0.25			
H2V1	18.64	1.64	0.14			
H2V2	0.00	0.00	0.00			
D	62.23	0.76	0.02	0.0228		
DV1	17.66	0.74	0.07			
DV2	0.80	0.75	0.15			
DH1	6.80	2.67	0.43			
DH2	35.02	3.95	0.42	0.0876	0.0527	
DH1V1	44.05	6.35	0.74	0.0554	0.0140	0.0359
DH1V2	58.43	2.91	0.38	0.0274	0.0965	
DH2V1	41.37	1.96	0.25	0.0634		
DH2V2	0.00	0.00	0.00			
Set 4						
VM	16.67	0.41	0.13			
HM	71.47	3.59	0.88	0.0149	0.0652	0.0220
HMVM	0.17	0.31	0.07			
DM	23.66	0.18	0.01			
DMVM	9.78	0.03	0.03			
DMHM	21.58	3.87	0.59		0.0554	0.0600
DMHMVM	75.56	4.12	0.57	0.0123	0.0482	0.0657

Table 4.12 (Continued)

Covariate set	Mean square			Prob > F[†]		
	HT	DBH	VOL	HT	DBH	VOL
Set 5						
VM1	30.54	0.91	0.26			
VM2	0.01	0.01	0.01			
HM1	51.26	1.98	0.77	0.0391		0.0328
HM2	22.20	1.79	0.16			
HM1VM1	2.77	0.44	0.11			
HM1VM2	0.29	0.33	0.11			
HM2VM1	30.23	1.97	0.16			
HM2VM2	0.00	0.00	0.00			
DM	22.97	0.18	0.01			
DMVM1	0.16	0.10	0.01			
DMVM2	0.38	0.47	0.08			
DMHM1	15.13	2.45	0.52			0.0778
DMHM2	14.00	2.45	0.22			
DMHM1VM1	30.79	2.42	0.30			
DMHM1VM2	58.26	2.23	0.34	0.0279		
DMHM2VM1	18.22	1.11	0.15			
DMHM2VM2	0.00	0.00	0.00			

[†]A blank implies that (Prob > F) > 0.10.

Table 4.13 Mean squares attributable to covariates independent of location of missing plants in relation to a standing plant and results of tests of significance of same for DATA LOBLOLLY PINE 2

Covariate set	Mean square			Prob > F†		
	HT	DBH	VOL	HT	DBH	VOL
Set 1						
C	134.07	3.91	0.60	0.0009	0.0543	0.0594
Set 6						
TM	92.72	2.70	0.61	0.0055	0.1100	0.0564
Set 7						
TM	92.72	2.70	0.61	0.0055		0.0564
TM*TM	52.55	1.27	0.22	0.0367		
Set 8						
T1M	125.58	3.06	0.61	0.0013	0.0887	0.0563
T2M	7.83	0.71	0.22			
T3M	2.47	0.09	0.03			
T4M	3.42	0.11	0.01			
Set 9						
S1M	98.57	2.06	0.38	0.0043		
S2M	19.51	1.01	0.26			
S3M	29.93	0.91	0.23			
S4M	0.38	0.03	0.02			
Set 10						
S1M	98.57	2.06	0.38	0.0043		
PT1M	34.84	1.71	0.45	0.0891		
T2M	0.00	0.00	0.00			
T3M	2.47	0.09	0.03			
T4M	3.42	0.11	0.01			
Set 11						
T1MT2M	133.41	3.62	0.76	0.0009	0.0641	0.0333
T3M	2.01	0.19	0.07			
T4M	3.39	0.11	0.01			
Set 12						
T1MT2M	133.41	3.62	0.76	0.0009	0.0641	0.0333
T3MT4M	0.83	0.01	0.01			
Set 13						
S1MS2M	115.16	2.71	0.53	0.0020		0.0764
S3M	29.71	0.89	0.22			
S4M	0.44	0.02	0.02			
Set 14						
S1MS2M	115.16	2.71	0.53	0.0020		0.0764
S3MS4M	17.45	0.76	0.21			
Set 15						
T1M	125.58	3.06	0.61	0.0013	0.0886	0.0562
MMT	0.04	0.01	0.01			
Set 16						
NT1M	111.92	3.73	0.93	0.0023	0.0602	0.0187
DM	29.40	0.34	0.03			

Table 4.13 (Continued)

Covariate set	Mean square			Prob > F†		
	HT	DBH	VOL	HT	DBH	VOL
Set 17						
NS1M	75.44	2.21	0.54	0.0124		0.0725
PT1M	42.64	2.01	0.52	0.0599		0.0782
DM	27.33	0.27	0.02			
Set 18						
NS1M	75.44	2.21	0.54	0.0124		0.0726
DM	31.78	0.38	0.03			
Set 19						
LTM	153.31	4.26	0.84	0.0004	0.0445	0.0255
Set 20						
LT1M	151.59	4.03	0.72	0.0004	0.0500	0.0382

†A blank implies that (Prob > F) > 0.10.

Table 4.14 Percent reduction in MSE and relative efficiency for adjustment of DATA LOBLOLLY PINE 2 using covariates in the 20 different sets

Set	Reduction in MSE (%)			Relative efficiency (%)		
	HT	DBH	VOL	HT	DBH	VOL
1	2.79	1.73	1.55	102.87	101.76	101.58
2	3.45	1.89	1.72	103.57	101.93	101.75
3	3.32	2.09	1.69	103.43	102.14	101.72
4	3.68	2.56	2.22	103.82	102.63	102.27
5	3.55	2.77	2.26	103.68	102.85	102.31
6	2.41	1.48	1.65	102.47	101.50	101.68
7	3.29	1.90	2.09	103.41	101.93	102.13
8	3.33	1.86	1.96	103.44	101.89	102.00
9	3.21	1.87	1.98	103.32	101.90	102.02
10	3.33	1.86	1.96	103.44	101.89	102.00
11	2.96	1.72	1.95	103.05	101.75	101.99
12	2.91	1.65	1.89	103.00	101.68	101.93
13	3.21	1.96	2.13	103.32	102.00	102.18
14	3.35	1.97	2.13	103.46	102.01	102.18
15	3.21	1.83	1.99	103.32	101.86	102.03
16	3.34	1.96	1.99	103.46	102.00	102.03
17	3.30	1.86	1.88	103.41	101.89	101.91
18	3.23	1.87	1.82	103.33	101.91	101.85
19	3.43	1.98	2.10	103.56	102.02	102.15
20	3.45	2.00	2.03	103.57	102.04	102.07

Table 4.15 Coefficient of variation (C.V.) expressed in percent derived from MSE for models with and without covariates for DATA LOBLOLLY PINE 2

Model	C.V. (%)		
	HT	DBH	VOL
Without covariates	46.43	51.33	125.13
With covariates (sets)			
1	45.78	50.89	124.76
2	45.62	50.84	124.05
3	45.65	50.79	124.07
4	45.57	50.67	123.74
5	45.60	50.62	123.71
6	45.86	50.95	124.10
7	45.66	50.84	123.82
8	45.65	50.85	123.90
9	45.68	50.85	123.89
10	45.65	50.85	123.90
11	45.74	50.89	123.91
12	45.75	50.91	123.95
13	45.68	50.83	123.79
14	45.64	50.82	123.79
15	45.68	50.86	123.88
16	45.64	50.83	123.88
17	45.65	50.85	123.95
18	45.67	50.85	123.99
19	45.62	50.81	123.81
20	45.62	50.82	123.86

HT, DBH, and VOL, respectively, and they differed little from those for the models with covariates. Therefore, in proceeding with the analysis of HT and VOL, only NT1M was used as a covariate. Only the covariate LTM was used in the analysis of DBH.

From Table 4.16 one sees that the F-values for treatments adjusted for NT1M were slightly larger than those for treatments unadjusted in the HT and DBH analyses. In like manner, the F-value for treatments adjusted for LTM was slightly larger than that for treatments unadjusted in the DBH analysis. In all of these cases, the F-values for adjusted treatments were significant at the P = 0.01 level. In Table 4.17 it is shown that the use of NT1M for HT and VOL, and LTM for DBH successfully decreased the MSE, especially in the case of the use of NT1M. The reductions in MSE were about 5%, 2%, and 3% for HT, DBH, and VOL, respectively.

The bias per missing plant surrounding a standing plant as measured by the regression coefficient of NT1M as a predictor of each of HT and VOL and that of LTM as a predictor of DBH is presented in Table 4.18. The negative bias was larger in magnitude for HT (-0.30 feet) with respect to VOL (-0.03 cubic feet) and DBH (-0.08 inches). The negative bias indicates that trees surrounded by missing trees did not reach the full development that standing trees with no surrounding missing trees would. The biases were significant at the P = 0.05 level, and their standard errors were about the same (0.10 feet for HT, 0.04 inches for DBH, and 0.01 cubic feet for VOL).

The regression coefficients and the total and means of each one of NT1M and LTM per treatment and for the entire experiment (Table 4.19)

Table 4.16 Unadjusted and adjusted mean squares using NT1M and LTM as the covariates and results of tests of significance of same for DATA LOBLOLLY PINE 2

Covariate	Variable	Source	d.f.	Mean square		F	
				Unadjusted	Adjusted	Unadjusted	Adjusted
NT1M	HT	Replications	3				
		Treatments	39	163.41	159.51	2.10**	2.16**
		Exp. error	114	77.90	73.89		
		Sampling error	2964				
	VOL	Replications	3				
		Treatments	39	1.48	1.47	2.39**	2.45**
		Exp. error	114	0.62	0.60		
		Sampling error	2964				
LTM	DBH	Replications	3				
		Treatments	39	10.79	10.77	3.06**	3.11**
		Exp. error	114	3.53	3.46		
		Sampling error	2964				

**Significant at P = 0.01.

Table 4.17 Percent reduction in MSE and relative efficiency for adjustment of DATA LOBLOLLY PINE 2 using the covariates NT1M and LTM

Covariate	Variable	Reduction in MSE (%)	Relative efficiency (%)
NT1M	HT	5.15	105.43
	VOL	3.23	103.33
LTM	DBH	1.98	102.02

Table 4.18 Simple regression coefficients of the response variables HT and VOL on the covariate NT1M and the response variable DBH on LTM and standard errors of same for DATA LOBLOLLY PINE 2

Covariate	Variable	Regression coefficient, b	Standard error of b	t for H_0: $\beta = 0$	Prob > |t|
NT1M	HT	-0.30050509	0.09856733	-3.04873	0.0023
	VOL	-0.02746341	0.01167503	-2.35232	0.0187
LTM	DBH	-0.07956601	0.03958374	-2.01007	0.0445

were used for computing the results presented in Tables 4.20, 4.21, 4.22, and 4.23.

Very small differences (less than 0.12 in magnitude for each of HT, DBH, and VOL) were observed between unadjusted and adjusted treatment means when the adjustment was to an equal average number of missing plants surrounding a standing plant (Table 4.20).

The estimated bias measured through the NT1M regression coefficient was larger in magnitude for HT (Table 4.21) than for VOL (Table 4.22) because of the large missing per-plant effect on HT. The total bias over all plants affected by missing plants for treatments ranged from -6.6111 to -18.0303 feet for HT and from -0.6042 to -1.6478 cubic feet for VOL. That for treatments measured by DBH regressed on LTM ranged from -1.7930 to -5.9878 inches (Table 4.23).

For the entire experiment, the percentage of missing plants was 15.19%. NT1M and LTM decreased the height, diameter, and volume of standing plants surrounded by missing plants by -472.9950 feet, -150.0512 inches and -43.2274 cubic feet, respectively.

The average biases for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment were -0.7911% for HT, -1.2972% for DBH, and -2.1517% for VOL.

From the results given for the loblolly pine experiments one may conclude that the effects due to missing plants were not constant for different response variables and environmental conditions. It was shown that HT was significantly affected by NT1M only in DATA LOBLOLLY PINE 2 whereas DBH was significantly affected by NT1M only in DATA LOBLOLLY PINE 1 and by LTM in DATA LOBLOLLY PINE 2. The variable VOL was significantly affected by NT1M but the per-plant bias was different

Table 4.19 Averages of the covariates NT1M and LTM for each treatment in DATA LOBLOLLY PINE 2

Treatment	n†	NT1M	Mean of NT1M	LTM	Mean of LTM
08009CC	82	49	0.5976	59.3598	0.7239
080090P	71	33	0.4648	49.1604	0.6924
08016CC	78	52	0.6667	32.6716	0.6283
080160P	90	42	0.4667	59.2200	0.6580
08035CC	65	30	0.4615	33.4880	0.5152
080350P	73	29	0.3973	46.0995	0.6315
08044CC	76	34	0.4474	37.2476	0.4901
080440P	79	30	0.3797	42.8101	0.5419
08058CC	70	40	0.5714	43.4630	0.6209
080580P	82	43	0.5244	56.4324	0.6882
08080CC	81	49	0.6049	50.4792	0.6232
080800P	92	45	0.4891	61.2904	0.6662
08103CC	99	60	0.6061	62.3997	0.6303
081030P	70	23	0.3286	29.6800	0.4240
08106CC	62	36	0.5806	32.6988	0.5274
081060P	85	49	0.5765	67.8640	0.7984
08117CC	69	27	0.3913	32.9820	0.4780
081170P	69	29	0.4203	31.3467	0.4543
08118CC	81	48	0.5926	44.0235	0.5435
081180P	93	44	0.4731	56.7486	0.6102
08120CC	72	40	0.5556	44.3160	0.6155
081200P	74	34	0.4595	39.7454	0.5371
08121CC	84	39	0.4643	51.3660	0.6115
081210P	88	33	0.3750	41.7736	0.4747
08122CC	87	49	0.5632	67.5816	0.7768
081220P	79	46	0.5823	51.1920	0.6480
08126CC	76	41	0.5395	42.7500	0.5625
081260P	87	23	0.2644	40.2201	0.4623
08127CC	83	49	0.5904	55.8092	0.6724
081270P	88	46	0.5227	55.7040	0.6330
08128CC	63	22	0.3492	22.5351	0.3577
081280P	72	31	0.4306	47.4912	0.6596
08129CC	75	42	0.5600	46.6875	0.6225
081290P	80	55	0.6875	75.2560	0.9407
08130CC	76	35	0.4605	41.9976	0.5526
081300P	74	40	0.5405	44.6516	0.6034
08140CC	64	28	0.4375	27.9104	0.4361
081400P	96	35	0.3646	47.6736	0.4966
08144CC	68	50	0.7353	55.9640	0.8230
081440P	68	44	0.6471	55.7804	0.8203
For the experiment	3121	1574	0.5043	1885.8706	0.6043

†Number of standing plants. The percentage of missing plants for the entire experiment was 15.19%.

Table 4.20 Unadjusted and adjusted treatment means when the adjustment was to an equal average number of missing plants surrounding a standing plant using NT1M and LTM as the covariates for DATA LOBLOLLY PINE 2

Treatment	n†	NT1M HT (feet) Unadjusted	NT1M HT (feet) Adjusted	NT1M VOL (cubic feet) Unadjusted	NT1M VOL (cubic feet) Adjusted	LTM DBH (inches) Unadjusted	LTM DBH (inches) Adjusted
08009CC	82	16.74	16.77	0.41	0.42	3.08	3.09
080090P	71	19.69	19.68	0.63	0.63	3.56	3.57
08016CC	78	16.14	16.18	0.50	0.50	2.96	2.96
080160P	90	17.66	17.65	0.56	0.56	3.61	3.61
08035CC	65	17.74	17.72	0.50	0.50	3.39	3.38
080350P	73	19.15	19.12	0.68	0.68	3.89	3.89
08044CC	76	18.57	18.56	0.57	0.57	3.51	3.50
080440P	79	19.08	19.04	0.65	0.65	3.88	3.88
08058CC	70	20.36	20.38	0.67	0.67	3.68	3.68
080580P	82	21.88	21.89	0.93	0.93	4.41	4.42
08080CC	81	19.03	19.06	0.58	0.58	3.47	3.47
080800P	92	21.06	21.05	0.88	0.88	4.24	4.24
08103CC	99	17.69	17.72	0.47	0.48	3.16	3.17
081030P	70	21.77	21.71	0.79	0.79	4.01	3.99
08106CC	62	18.21	18.24	0.47	0.48	3.26	3.26
081060P	85	19.67	19.69	0.75	0.75	4.02	4.04
08117CC	69	18.46	18.43	0.58	0.58	3.47	3.46
081170P	69	19.21	19.18	0.66	0.66	3.86	3.84
08118CC	81	18.34	18.37	0.52	0.52	3.35	3.34
081180P	93	19.07	19.06	0.63	0.63	3.70	3.70
08120CC	72	18.69	18.70	0.54	0.54	3.42	3.42
081200P	74	19.74	19.72	0.72	0.72	3.88	3.88
08121CC	84	19.25	19.24	0.62	0.62	3.66	3.66
081210P	88	19.51	19.48	0.65	0.64	3.71	3.70
08122CC	87	17.96	17.97	0.53	0.54	3.43	3.54
081220P	79	16.68	16.70	0.52	0.52	3.40	3.41

Table 4.20 (Continued)

Treatment	n†	NT1M				LTM	
		HT (feet)		VOL (cubic feet)		DBH (inches)	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
08126CC	76	17.71	17.72	0.48	0.48	3.17	3.17
081260P	87	18.78	18.70	0.67	0.66	3.89	3.87
08127CC	83	19.44	19.47	0.63	0.63	3.56	3.56
081270P	88	21.00	21.00	0.90	0.90	4.28	4.28
08128CC	63	19.83	19.78	0.64	0.64	3.79	3.77
081280P	72	20.10	20.07	0.82	0.82	4.32	4.32
08129CC	75	16.41	16.44	0.41	0.41	3.00	3.00
081290P	80	18.24	18.29	0.61	0.62	3.70	3.73
08130CC	76	21.12	21.11	0.81	0.81	3.99	3.99
081300P	74	22.08	22.09	0.92	0.92	4.23	4.23
08140CC	64	19.42	19.40	0.66	0.66	3.93	3.92
081400P	96	18.41	18.37	0.54	0.54	3.49	3.48
08144CC	68	17.58	17.65	0.47	0.48	3.23	3.24
081440P	68	19.47	19.52	0.68	0.68	3.76	3.78

†Number of standing plants.

Table 4.21 Estimated bias per treatment as measured by the covariate NT1M for the variable HT in DATA LOBLOLLY PINE 2

Treat-ment	Bias† Total	Bias† Average	Mean Unadjusted	Mean Adjusted	Total if no bias present
	feet				
08009CC	-14.7247	-0.1796	16.7390	16.9186	1691.8570
080090P	-9.9167	-0.1397	19.6901	19.8298	1982.9771
08016CC	-15.6263	-0.2003	16.1359	16.3362	1633.6237
080160P	-12.6212	-0.1402	17.6633	17.8035	1780.3536
08035CC	-9.0152	-0.1387	17.7364	17.8751	1787.5095
080350P	-8.7146	-0.1194	19.1534	19.2728	1927.2779
08044CC	-10.2172	-0.1344	18.5724	18.7068	1870.6836
080440P	-9.0152	-0.1141	19.0785	19.1926	1919.2616
08058CC	-12.0202	-0.1717	20.3571	20.5288	2052.8817
080580P	-12.9217	-0.1576	21.8841	22.0417	2204.1682
08080CC	-14.7247	-0.1818	19.0259	19.2077	1920.7687
080800P	-13.5227	-0.1470	21.0565	21.2035	2120.3486
08103CC	-18.0303	-0.1821	17.6869	17.8690	1786.9024
081030P	-6.9116	-0.0987	21.7657	21.8644	2186.4437
08106CC	-10.8182	-0.1745	18.2145	18.3890	1838.8987
081060P	-14.7243	-0.1732	19.6694	19.8426	1984.2632
08117CC	-8.1136	-0.1176	18.4638	18.5814	1858.1389
081170P	-8.7146	-0.1263	19.2087	19.3350	1933.4999
08118CC	-14.4242	-0.1781	18.3432	18.5213	1852.1277
081180P	-13.2222	-0.1422	19.0656	19.2078	1920.7774
08120CC	-12.0202	-0.1669	18.6892	18.8561	1885.6147
081200P	-10.2172	-0.1381	19.7351	19.8732	1987.3170
08121CC	-11.7197	-0.1395	19.2500	19.3895	1938.9520
081210P	-9.9167	-0.1127	19.5148	19.6275	1962.7489
08122CC	-14.7247	-0.1692	17.9568	18.1260	1812.6050
081220P	-13.8232	-0.1750	16.6810	16.8560	1685.5978
08126CC	-12.3207	-0.1621	17.7092	17.8713	1787.1315
081260P	-6.9116	-0.0794	18.7770	18.8564	1885.6444
08127CC	-14.7247	-0.1774	19.4410	19.6184	1961.8407
081270P	-13.8232	-0.1571	21.0011	21.1582	2115.8182
08128CC	-6.6111	-0.1049	19.8270	19.9319	1993.1938
081280P	-9.3157	-0.1294	20.0972	20.2266	2022.6584
08129CC	-12.6212	-0.1683	16.41321	16.5815	1658.1483
081290P	-16.5278	-0.2066	18.2395	18.4461	1844.6097
08130CC	-10.5177	-0.1384	21.1184	21.2568	2125.6791
081300P	-12.0202	-0.1624	22.0811	22.2435	2224.3535
08140CC	-8.4141	-0.1315	19.4156	19.5471	1954.7071
081400P	-10.5177	-0.1096	18.4115	18.5211	1852.1059
08144CC	-15.0253	-0.2210	17.5794	17.8004	1780.0360
081140P	-13.2222	-0.1944	19.4685	19.6629	1966.2944
For the experiment	-472.9950	-0.1516	19.0108	19.1624	76649.4096

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is -0.7911%.

Table 4.22 Estimated bias per treatment as measured by the covariate NT1M for the variable VOL in DATA LOBLOLLY PINE 2

Treat-ment	Bias†		Mean		Total
	Total	Average	Unadjusted	Adjusted	if no bias present
	----------	----------	Cubic feet	----------	----------
08009CC	-1.1457	-0.0164	0.4137	0.4301	43.0111
08009OP	-0.9063	-0.0128	0.6294	0.6422	64.2165
08016CC	-1.4281	-0.0183	0.4952	0.5135	51.3509
08016OP	-1.1535	-0.0128	0.5605	0.5733	57.3316
08035CC	-0.8239	-0.0127	0.5041	0.5168	51.6775
08035OP	-0.7964	-0.0109	0.6840	0.6949	69.4910
08044CC	-0.9338	-0.0123	0.5690	0.5813	58.1286
08044OP	-0.8239	-0.0104	0.6536	0.6640	66.4029
08058CC	-1.0985	-0.0157	0.6718	0.6875	68.7493
08058OP	-1.1809	-0.0144	0.9323	0.9467	94.6702
08080CC	-1.3457	-0.0166	0.5783	0.5949	59.4914
08080OP	-1.2359	-0.0134	0.8790	0.8924	89.2433
08103CC	-1.6478	-0.0166	0.4729	0.4895	48.9544
08103OP	-0.6317	-0.0090	0.7944	0.8034	80.3424
08106CC	-0.9887	-0.0159	0.4742	0.4901	49.0146
08106OP	-1.3457	-0.0158	0.7489	0.7647	76.4732
08117CC	-0.7415	-0.0107	0.5786	0.5893	58.9347
08117OP	-0.7964	-0.0115	0.6593	0.6708	67.0843
08118CC	-1.3182	-0.0163	0.5160	0.5323	53.2275
08118OP	-1.2084	-0.0130	0.6314	0.6444	64.4393
08120CC	-1.0985	-0.0153	0.5403	0.5556	55.5557
08120OP	-0.9338	-0.0126	0.7183	0.7309	73.0918
08121CC	-1.0711	-0.0128	0.6204	0.6332	63.3151
08121OP	-0.9063	-0.0103	0.6457	0.6560	65.5999
08122CC	-1.3457	-0.0155	0.5339	0.5494	54.9368
08122OP	-1.2633	-0.0160	0.5179	0.5339	53.3891
08126CC	-1.1260	-0.0148	0.4768	0.4916	49.1616
08126OP	-0.6317	-0.0073	0.6670	0.6743	67.4260
08127CC	-1.3457	-0.0162	0.6268	0.6430	64.3013
08127OP	-1.2633	-0.0144	0.8967	0.9111	91.1056
08128CC	-0.6042	-0.0096	0.6433	0.6529	65.2890
08128OP	-0.8514	-0.0118	0.8211	0.8329	83.2925
08129CC	-1.1535	-0.0154	0.4112	0.4266	42.6580
08129OP	-1.5105	-0.0189	0.6123	0.6312	63.1181
08130CC	-0.9612	-0.0126	0.8095	0.8221	82.2148
08130OP	-1.0985	-0.0148	0.9227	0.9375	93.7545
08140CC	-0.7690	-0.0120	0.6633	0.6753	67.5315
08140OP	-0.9612	-0.0100	0.5391	0.5491	54.9113
08144CC	-1.3732	-0.0202	0.4717	0.4919	49.1894
08144OP	-1.2084	-0.0178	0.6755	0.6933	69.3270
For the experi-ment	-43.2274	-0.0139	0.6321	0.6460	2583.8020

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is -2.1517%.

Table 4.23 Estimated bias per treatment as measured by the covariate LTM for the variable DBH in DATA LOBLOLLY PINE 2

Treat-ment	Bias† Total	Bias† Average	Mean Unadjusted	Mean Adjusted	Total if no bias present
	inches				
08009CC	-4.7230	-0.0576	3.0780	3.1356	313.5598
080090P	-3.9115	-0.0551	3.5606	3.6157	361.5692
08016CC	-2.5995	-0.0333	2.9577	2.9910	299.1028
080160P	-4.7119	-0.0524	3.6100	3.6624	366.2354
08035CC	-2.6645	-0.0410	3.3877	3.4287	342.8692
080350P	-3.6680	-0.0502	3.8865	3.9367	393.6746
08044CC	-2.9636	-0.0390	3.5092	3.5482	354.8195
080440P	-3.4062	-0.0431	3.8848	3.9279	392.7917
08058CC	-3.4582	-0.0494	3.6786	3.7280	372.8003
080580P	-4.4901	-0.0548	4.4134	4.4682	446.8157
08080CC	-4.0164	-0.0496	3.4691	3.5187	351.8686
080800P	-4.8766	-0.0530	4.2391	4.2921	429.2107
08103CC	-4.9649	-0.0502	3.1640	3.2142	321.4150
081030P	-2.3615	-0.0337	4.0056	4.0393	403.9336
08106CC	-2.6017	-0.0420	3.2645	3.3065	330.6463
081060P	-5.3997	-0.0635	4.0235	4.0870	408.7026
08117CC	-2.2642	-0.0380	3.4725	3.5105	351.0533
081170P	-2.4941	-0.0361	3.8551	3.8912	389.1247
08118CC	-3.5028	-0.0432	3.3494	3.3926	339.2644
081180P	-4.5153	-0.0486	3.7022	3.7508	375.0751
08120CC	-3.5260	-0.0490	3.4236	3.4726	347.2573
081200P	-3.1624	-0.0427	3.8824	3.9251	392.5135
08121CC	-4.0870	-0.0487	3.6595	3.7082	370.8155
081210P	-3.3238	-0.0378	3.7125	3.7503	375.0270
08122CC	-5.3772	-0.0618	3.4345	3.4963	349.6307
081220P	-4.0731	-0.0516	3.4025	3.4541	345.4059
08126CC	-3.4014	-0.0448	3.1724	3.2172	321.7156
081260P	-3.2002	-0.0368	3.8851	3.9219	392.1883
08127CC	-4.4405	-0.0535	3.5554	3.6089	360.8900
081270P	-4.4321	-0.0504	4.2807	4.3311	433.1065
08128CC	-1.7930	-0.0285	3.7857	3.8142	381.4161
081280P	-3.7787	-0.0525	4.3194	4.3719	437.1882
08129CC	-3.7147	-0.0495	3.0027	3.0522	305.2230
081290P	-5.9878	-0.0748	3.7000	3.7749	377.4848
08130CC	-3.3416	-0.0440	3.9908	4.0348	403.4768
081300P	-3.5527	-0.0480	4.2284	4.2764	427.6410
08140CC	-2.2207	-0.0347	3.9297	3.9644	396.4399
081400P	-3.7932	-0.0395	3.4854	3.5249	352.4912
08144CC	-4.4528	-0.0655	3.2250	3.2905	329.0483
081440P	-4.4382	-0.0653	3.7588	3.8241	382.4068
For the experiment	-150.0512	-0.0481	3.6598	3.7079	14831.5117

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is -1.2972%.

in both sets of data. Also, the percent reduction in MSE was larger in DATA LOBLOLLY PINE 1 than in DATA LOBLOLLY PINE 2.

The differences between DATA LOBLOLLY PINE 1 (positive bias) and DATA LOBLOLLY PINE 2 (negative bias) have also occurred because of the different loblolly pine families used in the experiments conducted at the two locations (Georgia and North Carolina). Location is known to affect the response variables as both data sets had different formulas for estimating the volume of a tree. Moreover, LOBLOLLY PINE 1 was a 15-year-old plantation whereas LOBLOLLY PINE 2 was a 10-year-old plantation at the time the data were recorded. The shape of the experimental unit may have also affected the bias. LOBLOLLY PINE 1 had a 5 x 5 tree square plot which allowed more within-family competition than among-family competition. LOBLOLLY PINE 2 had a one-row plot 25 trees long which would result in more competition among families than within a family. The differences in spacing between both data sets were considered irrelevant (eight and nine feet between trees for LOBLOLLY PINE 1 and LOBLOLLY PINE 2, respectively).

DATA OIL PALM. Of all of the response variables that were included in the analysis, B_1, Y_2, and MYB_2 were the only ones significantly affected by missing plants. Set 8 (with T1M, T2M, T3M, and T4M) and set 16 (with NT1M and DM) were of principal importance. In Table 4.24 is shown that the effect of T1M was clearly due to NT1M, whereas the diagonal effects were nonsignificant.

The comparison of these two models for each one of B_1, Y_2, and MYB_2 is made in Table 4.25. It was noted that the larger reductions in MSE were for MYB_2 in the amounts of 6.34% and 5.06% for sets 8 and 16, respectively; it then followed Y_2 with 3.59% (set 8) and 3.96% (set 6),

Table 4.24 Mean squares attributable to covariates in sets 8 and 16 and results of tests of significance of same for DATA OIL PALM

Covariate set	Mean square			Prob > F†		
	B_1	Y_2	MYB_2	B_1	Y_2	MYB_2
Set 8						
T1M	89.17	25647.27	159.57	0.0183	0.0162	0.0470
T2M	0.002	54.84	57.67			
T3M	3.97	2310.84	42.23			
T4M	6.45	16075.99	18.57			
Set 16						
NT1M	90.82	23761.38	165.26	0.0173	0.0206	0.0432
DM	16.12	125.89	45.12			

†A blank means that (Prob > F) > 0.10.

Table 4.25 Percent reduction in MSE, relative efficiency and C.V. for adjustment of DATA OIL PALM using the covariates in sets 8 and 16

Criterion of measurement	Model without covariates	Model with covariates	
		Set 8	Set 16
Reduction in MSE (%)			
B_1		2.64	2.04
Y_2		3.59	3.96
MYB_2		6.34	5.06
Relative efficiency (%)			
B_1	100	102.71	102.08
Y_2	100	103.73	104.12
MYB_2	100	106.77	105.33
C.V. (%)			
B_1	56.27	55.48	55.02
Y_2	49.48	49.03	49.25
MYB_2	55.27	54.27	55.07

and the last was B_1 with 2.64% (set 8) and 2.04% (set 6). So, both sets gave about the same results for the reduction in MSE. The covariate NT1M was used successfully to adjust the response variable giving larger and significant F-values (Table 4.26) and higher reductions in MSE of 5.40%, 4.75%, and 7.45% for B_1, Y_2, and MYB_2, respectively (Table 4.27).

The bias per missing plant surrounding a standing plant as measured by the regression coefficient of NT1M as a predictor of each of B_1, Y_2, and MYB_2 is presented in Table 4.28. The negative bias indicates a decrease of 0.44 $\pm$ 0.19 bunches, 7.49 $\pm$ 3.11 kg of bunches, and 0.59 $\pm$ 0.30 kg per bunch for each plant in each year if surrounded by missing plants with respect to plants with no surrounding missing plants. These biases were significant at the P = 0.05 level.

The regression coefficients and the total and mean of NT1M per treatment and for the entire experiment (Table 4.29) were used for computing the results shown in Tables 4.30, 4.31, 4.32, and 4.33.

In Table 4.30, the unadjusted and treatment means adjusted to an equal average number of missing plants surrounding a standing plant had larger differences for Y_2 (less than 7.97 kg of bunches in magnitude) than for B_1 (less than 0.48 bunches in magnitude), and MYB_2 (less than 1.06 kg per bunch in magnitude).

The estimated bias was larger in magnitude for Y_2 (Table 4.31) than for B_1 (Table 4.32) and MYB_2 (Table 4.33) because of the large missing per-plant effect on Y_2. The total bias for treatments ranged from -52.4357 to -883.9169 kg of bunches for Y_2, from -4.1361 to -69.7223 kg per bunch for MYB_2, and from -3.0918 to -52.1188 bunches for B_1.

For the entire experiment, the percentage of missing plants was 13.37% which had an overall effect on all of the standing plants

Table 4.26 Unadjusted and adjusted mean squares using NT1M as the covariate and results of tests of significance of same for DATA OIL PALM

Variable	Source	d.f.	Mean square		F	
			Unadjusted	Adjusted	Unadjusted	Adjusted
B_1	Replications	2				
	Treatments	7	454.20	433.45	54.53**	55.01**
	Exp. error	14	8.33	7.88		
	Sampling error	475				
Y_2	Replications	2				
	Treatments	7	31146.82	33372.43	9.93**	1.17**
	Exp. error	14	3135.55	2986.40		
	Sampling error	475				
MYB_2	Replications	2				
	Treatments	7	581.85	581.38	12.53**	13.52**
	Exp. error	14	46.45	42.99		
	Sampling error	475				

**Significant at P = 0.01.

Table 4.27 Percent reduction in MSE and relative efficiency for adjustment of DATA OIL PALM using the covariate NT1M

Covariate	Variable	Reduction in MSE (%)	Relative efficiency (%)
NT1M	B_1	5.40	105.71
	Y_2	4.75	104.99
	MYB_2	7.45	108.05

Table 4.28 Simple regression coefficients of the response variables B_1, Y_2 and MYB_2 on the covariate NT1M and standard errors of same for DATA OIL PALM

Covariate	Variable	Regression coefficient, b	Standard Error of b	t for H_0: $\beta = 0$	Prob > |t|
NT1M	B_1	-0.44168479	0.18609347	-2.37346	0.0180
	Y_2	-7.49082102	3.10821318	-2.41001	0.0163
	MYB_2	-0.59086701	0.29660632	-1.99209	0.0469

Table 4.29 Averages of the covariate NT1M for each treatment in DATA OIL PALM

Treatment	$n^{\dagger}$	NT1M	Mean of NT1M
212 x 430	69	47	0.6812
137 x 430	71	7	0.0986
254 x 250	68	14	0.2059
430 x 250	44	16	0.3636
137 x 113	37	14	0.3784
113 x 268	69	118	1.7101
137 x 250	72	57	0.7917
212 x 268	69	50	0.7246
For the experiment	499	323	0.6473

†Number of standing plants. The percentage of missing plants for the entire experiment was 13.37%.

Table 4.30 Unadjusted and adjusted treatment means when the adjustment was to an equal average number of missing plants surrounding a standing plant using NT1M as the covariate for DATA OIL PALM

Treatment	n[†]	B1 (bunches)		Y2 (kg of bunches)		MYB2 (kg/bunch)	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
212 x 430	69	6.67	6.68	134.13	134.39	11.75	11.77
137 x 430	71	10.08	9.84	140.82	136.71	8.49	8.17
254 x 250	68	2.00	1.81	72.93	69.62	17.07	16.81
430 x 250	44	2.75	2.62	109.53	107.41	13.74	13.58
137 x 113	37	6.97	6.85	99.45	97.44	7.48	7.32
113 x 268	69	3.93	4.40	124.12	132.08	10.92	11.55
137 x 250	72	4.38	4.44	104.17	105.25	14.51	14.60
112 x 268	69	4.07	4.11	111.52	112.10	12.99	13.04

[†]Number of standing plants.

Table 4.31 Estimated bias per treatment as measured by the covariate NTIM for the variable Y_2 in DATA OIL PALM

Treatment	Bias†		Mean		Total
	Total	Average	Unadjusted	Adjusted	if no bias present
	kg of bunches				
212 x 430	-352.0686	-5.1024	134.1333	139.2357	10024.9735
137 x 430	-52.4357	-0.7385	140.8225	141.5610	10192.3943
254 x 250	-104.8715	-1.5422	72.9265	74.4687	5361.7484
430 x 250	-119.8531	-2.7239	109.5341	112.2580	8082.5785
137 x 113	-104.8715	-2.8344	99.4514	102.2858	7364.5751
113 x 268	-883.9169	-12.8104	124.1232	136.9336	9859.2184
137 x 250	-426.9768	-5.9302	104.1653	110.0955	7926.8784
212 x 268	-374.5411	-5.4281	111.5174	116.9455	8420.0782
For the experiment	-2419.5352	-4.8488	113.1681	118.0169	67977.7159

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is -4.1086%.

Table 4.32 Estimated bias per treatment as measured by the covariate NT1M for the variable B_1 in DATA OIL PALM

Treatment	Bias†		Mean		Total
	Total	Average	Unadjusted	Adjusted	if no bias present
	bunches				
212 x 430	-20.7592	-0.3009	6.6667	6.9676	501.6672
137 x 430	-3.0918	-0.0435	10.0845	10.1280	729.2160
254 x 250	-6.1836	-0.0909	2.000	2.0909	150.5448
430 x 250	-7.0670	-0.1606	2.7500	2.9106	209.5632
137 x 113	-6.1836	-0.1671	6.9730	7.1401	514.0872
113 x 268	-52.1188	-0.7553	3.9275	4.6828	337.1616
137 x 250	-25.1760	-0.3497	4.3750	4.7247	340.1784
212 x 268	-22.0842	-0.3201	4.0725	4.3926	320.6598
For the experiment	-142.6642	-0.2859	5.1263	5.4122	3117.4272

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is -5.2825%.

Table 4.33 Estimated bias per treatment as measured by the covariate NTIM for the variable MYB_2 in DATA OIL PALM

Treatment	Bias[†]		Mean		Total
	Total	Average	Unadjusted	Adjusted	if no bias present
	kg per bunch				
212 x 430	-27.7707	-0.4025	11.7511	12.1536	875.0574
137 x 430	-4.1361	-0.0583	8.4922	8.5505	615.6327
254 x 250	-8.2721	-0.1216	17.0679	17.1895	1237.6475
430 x 250	-9.4539	-0.2149	13.7447	13.9596	1005.0884
137 x 113	-8.2721	-0.2236	7.4784	7.7020	554.5419
113 x 268	-69.7223	-1.0105	10.9218	11.9323	859.1233
137 x 250	-33.6794	-0.4678	14.5111	14.9789	1078.4786
212 x 268	-29.5434	-0.4282	12.9922	13.4204	966.2662
For the experiment	-190.8500	-0.3825	12.3261	12.7086	7320.1335

[†]The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is -3.0098%.

surrounded by missing plants of -2419.5352 kg of bunches for Y_2, -190.8500 kg per bunch for MYB_2, and -142.6642 bunches for B_1. The average biases for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment were -4.1086% for Y_2, -3.0098% for MYB_2, and -5.2825% for B_1.

For DATA OIL PALM, the estimated negative bias was a result of the particular response variable in the experiment and the environmental condition during the year of harvest. Moreover, the wide spacing between plants (10 m) within and between treatments may have lessened the competitive effects. This seems reasonable because the bias was only significantly estimated by NT1M in the first two years of harvest (about four years after the experiment was established).

DATA PEACH. The design of an experiment, data from which are used to evaluate the estimation of adjustment procedures proposed in this thesis, should be relatively simple and straightforward.

This experiment had a very complex design. There were a number of factors involved and additional complexities were brought into the analysis when the covariates were included. To reduce the difficulties in analysis, a separate analysis was conducted for each variety (Candor and Redskin) in each year (1969, 1970, 1971, and 1972). Even with this simplification, negative predicted values for some missing observations (missing trees) were obtained using the models both with and without covariates. The logarithmic function was used to transform the response variables (TOH, NOFR, and TRAR) in order to obviate this problem but it was not successful. Nevertheless, some results are presented for variety Candor and variables TRAR (in 1970 and 1971) and TOH (in 1971 and 1972) for which the covariate TM was significant. The significance of TM was an indication that missing plants could have affected the response variables but the negative signs of the predicted values imply that perhaps the presence of the many design variables in the model caused the adjustment procedure some difficulties.

Since spacing was a factor studied in this experiment, the interaction of TM with spacing was considered in order to determine the necessity of defining the covariates for each level of spacing. However, this interaction was not significant. Besides, there were more than four diagonal locations and the absence of some vertical or horizontal locations happened for some standing plants. In this case, a diagonal location was taken as a vertical or horizontal location, depending upon the closeness of the diagonal location to what should be the vertical or horizontal location. An example is presented in Figure 4.1 where spacing between plants within rows is 20 feet (rows 33 and 34) and 12 feet (row 32); d_5 and d_6 are two additional diagonal locations; h_2 is 20 feet from p, whereas d_5 is 20.4 feet from p. Then, one could take (as was done here) $h_1 = d_5$; d_6 would be left as another diagonal location or may be added to d_5, i.e., $h_1 = d_5 + d_6$. Also, the effects of missing plants along rows with 12 feet spacing between plants were considered the same as those along rows with 20 feet spacing between plants. This was a consequence of the lack of significance of the interaction of TM with spacing. Therefore, the covariates for these data followed the definitions given previously for the other data sets with the addition of some more diagonal locations for some plants when defining D and DM.

Discussion concerning covariates will be focused on sets 8 and 16 which were the more important for studying the effects of missing plants. All other covariates for these data were considered irrelevant.

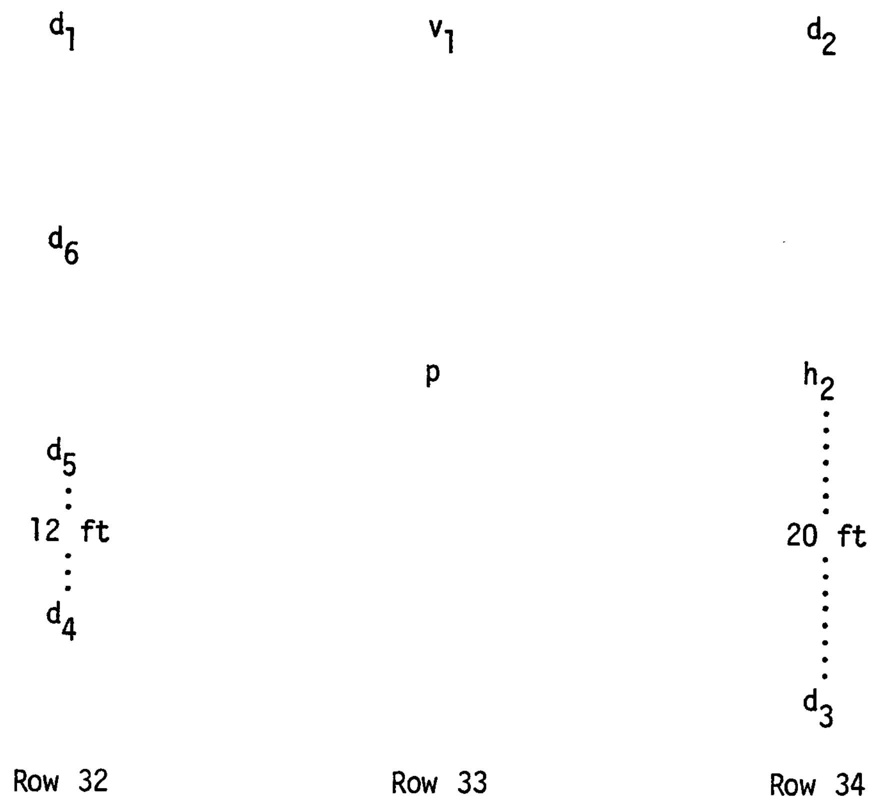

Figure 4.1 Locations of missing plants surrounding a standing plant, p, for a planting pattern with different spacing levels, 12 and 20 feet, between plants within a row for DATA PEACH

The variation due to missing plants was explained more successfully by set 16 than set 8. This is shown in Table 4.34 where NT1M was significant for both TRAR and TOH. The further study of T1M = NT1M + DM indicated that the diagonal effects given by DM were not significant. It appears that DM has interfered with the interpretation of T1M for TRAR (1970), TOH (1971) and TOH (1972) for which T1M was not significant; this is more noticeable for TOH (1971) where T1M was nonsignificant with probability larger than 0.50. For TOH (1971), T2M had the largest significant effect; the significance of T4M indicates that covariates like T5M, T6M, ..., TℓM (for some $\ell > 4$) should be included in the analysis of covariance. However, this was not done because the difference between both sets was relatively insignificant as it is shown in Table 4.35. In this table, the values for the reductions in MSE were 14.61% (set 8) and 13.83% (set 16), those for the relative efficiencies were 117.11% (set 8) and 116.05% (set 16), and those for the C.V.'s were 19.20% (set 8) and 19.29% (set 16). Therefore, an adjustment with NT1M as the only covariate in the model seemed appropriate for studying the effects of missing plants.

Although it was expected that all terms adjusted for NT1M in the analysis of variance would have diminished mean squares, this was not generally the case (Tables 4.36 and 4.37). The corresponding F-values did not change considerably and the same significance of the unadjusted and adjusted factorial effects was obtained, except for RP*PRE*S (TRAR, 1970), NP (TOH, 1971) and RP*PRE*NP, RP*PRE*S*NP (TOH, 1972). The difference between spacing levels had the largest contribution to the total variation for both TRAR and TOH; however, it was not significant for TOH (1972).

Table 4.34 Mean squares attributed to covariates in sets 8 and 16 in different years and results of tests of significance of same for variety Candor, DATA PEACH

Covariate set	Mean square		Prob > F†	
	TOH	TRAR	TOH	TRAR
	Year 1970			
Set 8				
T1M	400.67	135124.05		0.0740
T2M	2.27	22646.34		
T3M	616.69	116119.51		0.0973
T4M	72.55	132446.54		0.0766
Set 16				
NT1M	52.60	189070.23		0.0369
DM	375.91	20358.31		
	Year 1971			
Set 8				
T1M	172.50	337959.04	0.5119	0.0367
T2M	5112.90	31717.52	0.0005	
T3M	18.94	36005.21		
T4M	1709.97	91189.74	0.0405	
Set 16				
NT1M	2720.06	519012.66	0.0127	0.0091
DM	237.94	62915.33		
	Year 1972			
Set 8				
T1M	692.91	68636.18	0.0641	
T2M	35.83	278520.42		
T3M	1.80	18421.70		
T4M	103.78	337992.04		0.0884
Set 16				
NT1M	985.19	436296.48	0.0245	0.0539
DM	208.80	317.20		

†A blank implies that (Prob > F) > 0.10.

Table 4.35 Percent reduction in MSE, relative efficiency, and C.V. for adjustment of DATA PEACH using the covariates in sets 8 and 16

Criterion of measurement	Model without covariates	Model with covariates Set 8	Set 16
Reduction in MSE (%)†			
TRAR (1970)		0.76	9.95
TRAR (1971)		5.23	6.78
TOH (1971)		14.61	13.83
TOH (1972)		1.25	4.91
Relative efficiency (%)			
TRAR (1970)	100	100.77	111.05
TRAR (1971)	100	105.52	107.27
TOH (1971)	100	117.11	116.05
TOH (1972)	100	101.26	105.16
C.V. (%)			
TRAR (1970)	19.39	19.32	18.40
TRAR (1971)	19.82	19.29	19.13
TOH (1971)	20.78	19.20	19.29
TOH (1972)	23.99	23.39	23.39

†Based on plot-to-plant variation, E(f).

Table 4.36 Unadjusted and adjusted mean squares using NT1M as the covariate and results of tests of significance of same for TRAR, variety Candor, DATA PEACH

Year	Source	d.f.	Mean square		F	
			Unadjusted	Adjusted	Unadjusted	Adjusted
1970	REP	2				
	NP	3	1486829.40	1476856.69	8.84*	8.88*
	E(a)	6	168232.49	166307.86		
	S	1	7104142.34	7091580.48	46.42*	43.07*
	E(b)	2	153048.08	164664.24		
	S*NP	3	53632.86	51156.28	0.53	0.51
	E(c)	6	100840.78	101027.67		
	PRE	1	815.70	1347.59	0.002	0.004
	PRE*S	1	25707.44	21944.95	0.07	0.06
	E(d)	4	345761.15	354725.60		
	PRE*NP	3	52624.40	53088.07	0.86	0.89
	PRE*S*NP	3	8147.77	9092.39	0.13	0.15
	E(e)	12	61352.72	59922.33		
	RP	3	354297.49	342831.19	6.35*	5.90*
	RP*NP	9	33595.22	34044.35	0.60	0.59
	RP*S	3	135938.68	140519.71	2.43	2.42
	RP*S*NP	9	62894.95	63056.64	1.13	1.09
	RP*PRE	3	87766.56	90726.99	1.57	1.56
	RP*PRE*NP	9	70142.70	70422.61	1.26	1.29
	RP*PRE*S	3	156520.90	146803.82	2.80*	2.53
	RP*PRE*S*NP	9	84832.38	84671.15	1.52	1.46
	E(f)	69	55827.19	58063.08		

Table 4.36 (Continued)

Year	Source	d.f.	Mean square		F	
			Unadjusted	Adjusted	Unadjusted	Adjusted
1971	REP	2				
	NP	3	1235327.60	1292067.38	9.97*	9.47*
	E(a)	6	123853.23	136424.60		
	S	1	15602105.12	15783977.04	23.35*	25.51*
	E(b)	2	668255.41	618787.78		
	S*NP	3	110145.88	116881.13	0.46	0.47
	E(c)	6	238173.82	248072.47		
	PRE	1	23270.28	16484.60	0.06	0.03
	PRE*S	1	132531.26	111268.15	0.31	0.20
	E(d)	4	421728.02	555896.34		
	PRE*NP	3	28624.40	34090.28	0.25	0.40
	PRE*S*NP	3	162352.13	158038.90	1.42	1.87
	E(e)	12	114694.34	84396.52		
	RP	3	767126.63	674846.45	7.87*	6.88*
	RP*NP	9	74762.21	68281.17	0.77	0.70
	RP*S	3	109939.26	101725.34	1.13	1.04
	RP*S*NP	9	94129.25	100676.17	0.97	1.03
	RP*PRE	3	129148.15	146448.44	1.32	1.49
	RP*PRE*NP	9	79349.26	78440.78	0.81	0.80
	RP*PRE*S	3	35622.34	15245.90	0.37	0.16
	RP*PRE*S*NP	7	127962.18	145824.68	1.31	1.49
	E(f)	60	97481.66	98103.87		

*Significant at P = 0.05. All other terms without (*) are non-significant at P = 0.05.

Table 4.37 Unadjusted and adjusted mean squares using NT1M as the covariate and results of tests of significance of same for TOH variety Candor, DATA PEACH

Year	Source	d.f.	Mean square		F	
			Unadjusted	Adjusted	Unadjusted	Adjusted
1971	REP	2				
	NP	3	16819.43	17029.24	4.60	4.76*
	E(a)	6	3658.40	3576.94		
	S	1	114911.43	115329.43	50.95*	54.11*
	E(b)	2	2255.21	2131.48		
	S*NP	3	164.61	152.48	0.13	0.12
	E(c)	6	1285.64	1275.91		
	PRE	1	76.79	68.35	0.03	0.02
	PRE*S	1	494.68	468.00	0.17	0.14
	E(d)	4	2968.63	3387.81		
	PRE*NP	3	1603.55	1743.59	1.86	1.95
	PRE*S*NP	3	252.42	244.62	0.29	0.27
	E(e)	12	863.33	892.31		
	RP	3	3025.30	2500.16	4.65*	3.70*
	RP*NP	9	622.25	821.59	0.96	1.21
	RP*S	3	1051.10	1084.04	1.62	1.60
	RP*S*NP	9	839.33	850.13	1.29	1.26
	RP*PRE	3	560.43	681.86	0.86	1.01
	RP*PRE*NP	9	556.22	570.43	0.86	0.84
	RP*PRE*S	3	1392.26	1307.48	2.14	1.93
	RP*PRE*S*NP	7	656.80	610.49	1.01	0.90
	E(f)	60	650.24	676.59		

Table 4.37 (Continued)

Year	Source	d.f.	Mean square		F	
			Unadjusted	Adjusted	Unadjusted	Adjusted
1972	REP	2				
	NP	3	5380.16	5413.82	2.64	2.69
	E(a)	5	2037.10	2011.95		
	S	1	31919.82	31945.46	9.74	9.76
	E(b)	2	3275.97	3273.62		
	S*NP	3	343.28	346.54	0.86	0.85
	E(c)	5	399.29	406.12		
	PRE	1	71.24	71.06	0.06	0.05
	PRE*S	1	1072.90	1033.84	0.84	0.78
	E(d)	4	1281.19	1332.48		
	PRE*NP	3	318.50	285.62	0.88	0.81
	PRE*S*NP	3	210.72	198.26	0.58	0.56
	E(e)	9	363.66	354.28		
	RP	3	2254.93	2291.96	7.68*	6.97*
	RP*NP	9	705.20	737.19	2.40*	2.24*
	RP*S	3	850.72	841.65	2.90	2.56
	RP*S*NP	9	445.85	442.03	1.52	1.34
	RP*PRE	3	119.05	104.66	0.41	0.32
	RP*PRE*NP	9	666.10	652.48	2.27*	1.98
	RP*PRE*S	3	250.18	240.25	0.85	0.73
	RP*PRE*S*NP	6	768.27	742.79	2.62*	2.26
	E(f)	28	293.71	348.88		

*Significant at P = 0.05. All other terms without (*) are non-significant at P = 0.05.

The effect of NT1M on each of the error terms was different among and within years; this is clearly observed in Table 4.38 where the reductions for the error terms did not follow any definitive trend. The extreme effects were found for TRAR (1971) which had reductions of -31.82% for E(d) and 26.42% for E(e).

The bias per missing plant surrounding a standing plant as measured by the regression coefficient of NT1M as a predictor of each of TRAR (1970), TRAR (1971), TOH (1971), and TOH (1972) is presented in Table 4.39. The biases for TRAR (70.08 $\pm$ 33.16 and 116.87 $\pm$ 44.04 square inches in 1970 and 1971, respectively) were larger than the biases for TOH (8.46 $\pm$ 3.34 and 6.22 $\pm$ 2.72 pounds of fruit per tree in 1971 and 1972, respectively).

The regression coefficients and the total and mean of NT1M per treatment and for the entire experiment (Table 4.40) were used in the computation of the results presented in Tables 4.41, 4.42, 4.43, 4.44 and 4.45.

The differences between unadjusted and adjusted treatment means when the adjustment was to an equal average number of missing plants surrounding a standing plant were less than 27.10 square inches for TRAR and less than 2.60 pounds of fruit for TOH (Table 4.41).

The total bias over all plants affected by missing plants for factorial effect ranged from 1401.5584 to 5676.3117 square inches for TRAR (1970), from 2571.1053 to 12738.6579 square inches for TRAR (1971), from 186.1316 to 922.1974 pounds of fruit for TOH (1971), and from 111.8824 to 745.8824 pounds of fruit for TOH (1972).

For the entire experiment, the percentages of missing plants were 22.66% in 1970, 27.60% in 1971 and 47.14% in 1972. These percentages

Table 4.38 Percent reduction in MSE and relative efficiency for adjustment of DATA PEACH, variety Candor, using the covariate NT1M

Criterion of measurement	TRAR (1970)	TRAR (1971)	TOH (1971)	TOH (1972)
Reduction in MSE (%)				
E(a)	1.15	-10.14	2.23	1.23
E(b)	-7.58	7.40	5.32	0.07
E(c)	-0.18	-4.16	0.75	-1.71
E(d)	-2.60	-31.82	-14.12	-4.00
E(e)	2.33	26.42	-3.36	2.58
E(f)	-4.00	-0.63	-4.05	-11.97
Relative efficiency (%)				
E(a)	101.16	90.79	102.28	101.25
E(b)	92.95	107.99	105.62	100.07
E(c)	99.82	96.01	100.76	98.32
E(d)	97.47	75.86	87.63	96.15
E(e)	102.39	135.90	96.75	102.65
E(f)	96.15	99.37	96.11	89.31

Table 4.39 Simple regression coefficients of the response variables TRAR (1970), TRAR (1971), TOH (1971) and TOH (1972) on the covariate NT1M and standard errors of same for DATA PEACH

Covariate	Variable	Regression coefficient, b	Standard error of b	t for H_0: $\beta = 0$	Prob > |t|
NT1M	TRAR (1970)	70.07792208	33.16074231	2.11328	0.0365
	TRAR (1971)	116.86842105	44.04030699	2.65367	0.0090
	TOH (1971)	8.46052632	3.33947682	2.53349	0.0126
	TOH (1972)	6.21568627	2.71627761	2.28831	0.0246

Table 4.40 Averages of the covariate NT1M for each factorial effect in connection with each of the response variables for DATA PEACH

Factorial effect		TRAR (1970)			TRAR (1971) and TOH (1971)			TOH (1972)		
		n[†]	NT1M	Mean of NT1M	n	NT1M	Mean of NT1M	n	NT1M	Mean of NT1M
NP:	11	75	43	0.5733	66	56	0.8485	46	56	1.2174
	12	78	44	0.5641	75	57	0.7600	66	71	1.0758
	21	70	35	0.5000	65	44	0.6769	42	44	1.0476
	22	74	40	0.5405	72	50	0.6944	49	49	1.0000
S:	1	143	81	0.5664	135	103	0.7630	98	107	1.0918
	2	154	81	0.5260	143	104	0.7273	105	113	1.0762
PRE:	1	144	81	0.5625	139	98	0.7050	97	100	1.0309
	2	153	81	0.5294	139	109	0.7842	106	120	1.1321
RP:	11	65	36	0.5538	58	46	0.7931	43	44	1.0233
	12	94	49	0.5213	92	65	0.7065	84	100	1.1905
	21	50	20	0.4000	43	22	0.5116	27	18	0.6667
	22	88	57	0.6477	85	74	0.8706	49	58	1.1837
For the experiment		297	162	0.5455	278	207	0.7446	203	220	1.0837

†Number of standing plants. The percentages of missing plants for the entire experiment were 22.66%, 27.60% and 47.14% in 1970, 1971 and 1972, respectively.

Table 4.41 Unadjusted and adjusted factorial effect means when the adjustment was to an equal average number of missing plants surrounding a standing plant using NT1M as the covariate for variety Candor, DATA PEACH

Factorial effect		n†	Unadjusted	Adjusted	n	Unadjusted	Adjusted
			TRAR (1970)			TRAR (1971)	
			square inches				
NP:	11	75	1155.68	1153.73	66	1478.41	1466.27
	12	78	1057.71	1056.40	75	1437.13	1435.33
	21	70	1358.23	1361.41	65	1681.67	1688.83
	22	74	1318.97	1319.33	72	1710.70	1717.28
S:	1	143	1054.22	1052.75	135	1332.90	1330.76
	2	154	1370.81	1372.17	143	1804.49	1806.52
PRE:	1	144	1230.81	1229.62	139	1575.40	1580.03
	2	153	1206.67	1207.80	139	1575.56	1570.94
RP:	11	65	1203.97	1203.38	58	1512.21	1506.54
	12	94	1157.65	1159.34	92	1498.22	1503.68
	21	50	1189.60	1199.79	43	1527.82	1554.91
	22	88	1310.24	1303.07	85	1727.79	1713.06
			TOH (1971)			TOH (1972)	
			pounds of fruit				
NP:	11	66	112.89	112.02	46	55.02	54.19
	12	75	106.92	106.79	66	73.52	73.56
	21	65	131.68	132.20	42	75.07	75.30
	22	72	139.49	139.97	49	81.00	81.52
S:	1	135	102.35	102.19	98	58.67	58.62
	2	143	141.92	142.07	105	83.38	83.43
PRE:	1	139	123.15	123.49	97	70.70	71.03
	2	139	122.26	121.92	106	72.14	71.84
RP:	11	58	128.21	127.80	43	71.56	71.93
	12	92	115.38	115.77	84	63.83	63.17
	21	43	129.46	131.42	27	84.22	86.81
	22	85	124.13	123.06	49	77.39	76.77

†Number of standing plants.

Table 4.42 Estimated bias per factorial effect as measured by the covariate NT1M for the variable TRAR (1970) in DATA PEACH

Factorial effect		Bias†		Mean		Total
		Total	Average	Unadjusted	Adjusted	if no bias present
		square inches				
NP:	11	3013.3506	40.1780	1155.6800	1115.5020	107088.1912
	12	3083.4286	39.5311	1057.7051	1018.1740	97744.7006
	21	2452.7273	35.0390	1358.2286	1323.1896	127026.2053
	22	2803.1169	37.8800	1318.9865	1281.1065	122986.2280
S:	1	5676.3117	39.6945	1054.2238	1014.5293	194789.6280
	2	5676.3117	36.8592	1370.8052	1333.9460	256117.6384
PRE:	1	5676.3117	39.4188	1230.8125	1191.3937	228747.5844
	2	5676.3117	37.1001	1206.6732	1169.5731	224558.0397
RP:	11	2522.8052	38.8124	1203.9692	1165.1568	111855.0540
	12	3433.8182	36.5300	1157.6489	1121.1189	107627.4163
	21	1401.5584	28.0312	1189.6000	1161.5688	111510.6078
	22	3994.4416	45.3914	1310.2686	1264.8772	121428.2130
For the experiment		11352.6234	38.2243	1218.3771	1180.1528	453178.6671

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is 3.2389%.

Table 4.43 Estimated bias per factorial effect as measured by the covariate NTIM for the variable TRAR (1971) in DATA PEACH

Factorial effect		Bias†		Mean		Total
		Total	Average	Unadjusted	Adjusted	if no bias present
		square inches				
NP:	11	6544.6316	99.1611	1478.4091	1379.2480	132407.8095
	12	6661.5000	88.8200	1437.1333	1348.3133	129438.0768
	21	5142.2105	79.1109	1681.6667	1602.5558	153845.3538
	22	5843.4211	81.1586	1710.7013	1629.5427	156436.0967
S:	1	12037.4474	89.1663	1332.9037	1243.7374	238797.5853
	2	12154.3158	84.9952	1804.4895	1719.4943	330142.9027
PRE:	1	11453.1053	82.3964	1575.4929	1493.0965	286674.5202
	2	12738.6579	91.6450	1575.5612	1483.9162	284911.9064
RP:	11	5375.9474	92.6887	1512.2069	1419.5182	136273.7426
	12	7596.4474	82.5701	1498.2188	1415.6487	135902.2771
	21	2571.1053	59.7931	1527.8205	1468.0274	140930.6260
	22	8648.2632	101.7443	1727.7882	1626.0439	156100.2170
For the experiment		24191.7632	87.0207	1575.4820	1488.4613	571569.1274

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is 5.8464%.

Table 4.44 Estimated bias per factorial effect as measured by the covariate NT1M for the variable TOH (1971) in DATA PEACH

Factorial effect		Bias†		Mean		Total
		Total	Average	Unadjusted	Adjusted	if no bias present
		pounds of fruit				
NP:	11	473.7895	7.1786	112.8939	105.7153	10148.6661
	12	482.2500	6.4300	106.9200	100.4900	9647.0400
	21	372.2632	5.7271	131.6833	125.9562	12091.7928
	22	423.0263	5.8754	139.4935	133.6181	12827.3409
S:	1	871.4342	6.4551	102.3481	95.8930	18411.4621
	2	879.8947	6.1531	141.9231	135.7700	26067.8381
PRE:	1	829.1316	5.9650	123.1511	117.1861	22499.7359
	2	922.1974	6.6345	122.2590	115.6245	22199.9014
RP:	11	389.1842	6.7101	128.2069	121.4968	11663.6954
	12	549.9342	5.9775	115.3750	109.3975	10502.1556
	21	186.1316	4.3286	129.4615	125.1329	12012.7544
	22	626.0789	7.3656	124.1294	116.7638	11209.3215
For the experiment		1751.3289	6.2997	122.7050	116.4053	44899.6181

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is 5.4119%.

Table 4.45 Estimated bias per factorial effect as measured by the covariate NT1M for the variable TOH (1972) in DATA PEACH

Factorial effect		Bias†		Mean		Total
		Total	Average	Unadjusted	Adjusted	if no bias present
		pounds of fruit				
NP:	11	348.0784	7.5669	55.0217	47.4848	4558.5386
	12	441.3137	6.6866	73.5152	66.8286	6415.5483
	21	273.4902	6.5117	75.0714	68.5597	6581.7340
	22	304.5686	6.2157	81.0000	74.7843	7179.2941
S:	1	665.0784	6.7865	58.6735	51.8870	9962.3012
	2	702.3725	6.6893	83.3810	76.6917	14724.8136
PRE:	1	621.5686	6.4079	70.7010	64.2931	12344.2706
	2	745.8824	7.0366	72.1415	65.1049	12500.1358
RP:	11	273.4902	6.3602	71.5581	65.1979	6258.9948
	12	621.5686	7.3996	63.8333	56.4337	5417.6327
	21	111.8824	4.1438	84.2222	80.0784	7687.5273
	22	360.5098	7.3573	77.3878	70.0305	6722.9239
For the experiment		1367.4510	6.7362	71.4532	64.7170	24851.3235

†The average bias for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment is 10.4087%.

had overall effects on all of the standing plants surrounded by missing plants of 11352.6234 and 24191.7632 square inches for TRAR (1970) and TRAR (1971), respectively, and 1751.3289 and 1367.4510 pounds of fruit for TOH (1971) and TOH (1972), respectively. The average biases for the entire experiment expressed as a percentage of the adjusted mean for the entire experiment were 3.2389% for TRAR (1970), 5.8464% for TRAR (1971), 5.4119% for TOH (1971), and 10.4087% for TOH (1972).

For DATA PEACH, the experiment was conducted at the same experimental site through all the years. The data were first recorded about three years after the experiment was established. The estimated positive bias was dependent upon the response variable and the among years environmental variation. There were not detectable significant missing plant effects for the first year of harvest whereas they were detected in data for the following three years. Spacing did not seem to interfere with the estimation of the missing plant effects as did the many design variables.

The results of the peach experiment indicated the usefulness of the covariates for adjusting the means of some of the response variables, even though this experiment had a complex design and a high percentage (47.14%) of missing plants for the fourth harvest.

To proceed with the analysis of the data after adjusting for the missing plant effects, a covariance analysis of plot totals adjusted for the missing plant effects using stand differences as a covariate will be discussed.

The variable yield of bunches (Y_2) in DATA OIL PALM was adjusted for stand differences after adjusting for the missing plant effects. Two approaches were considered for this adjustment. The first was the covariance analysis of Y_2 on stand after adjustment for the missing plant effects has been made using the per-plant covariance analysis approach described in this thesis. The second was the covariance analysis of Y_2 on stand after adjustment for the missing plant effects has been performed using an estimator of the missing plant effects based upon the covariance analysis of plot totals and the covariate NT1M per plot. The second approach is intended to be used if computing facilities are inadequate for carrying out the first approach. For this particular case, the second approach was only an approximate analysis because standing plants with no yield were excluded from the analysis.

The results for the discussion of these two approaches are presented in Tables 4.46 and 4.47. There was a low relative efficiency (based upon the average mean square error for pairwise treatment mean comparisons) of 96.81% for the adjustment of yields based only upon stand. This was due to the large contribution of stand variation to the yield variation. Although, the MSE was reduced by 63.05%. This stand variation did not affect the effectiveness of both approaches very much for which the relative efficiencies were 114.69% for the first approach and 116.14% for the second approach. The reductions in MSE were 68.81% for the first approach and 69.20% for the second approach. This indicated clearly the usefulness of adjusting yields for the missing plant effects before adjusting yields for stand irregularities.

The bias estimated from the plot data analysis was -5.35 ± 5.52 kg of bunches whereas the bias estimated from the per-plant data analysis was -7.49 ± 3.11 kg of bunches. The close similarity of the results allows to use the second approach as an approximation for the first approach.

Table 4.46 Mean square error obtained in the covariance models with NT1M and stand as the covariates for the variable Y_2 in DATA OIL PALM

Covariates in model	MSE for	
	Overall F-test	Pairwise treatment comparison
None	56400.75830	56400.75830
NT1M†		
per-plant data	57307.03622	57307.03622
per-plot data	56644.02798	56644.02798
NT1M (per-plot data)††	56645.47349	70380.91219†††
Stand	20842.36490	58259.88028†††
Stand after adjusting plot totals for NT1M:		
per-plant data	17875.11157	49965.62840†††
per-plot data	17447.53028	48770.42647†††
NT1M and stand	17419.34177	52666.46161†††

†The MSE's were estimated from the analysis of variance of plot totals adjusted for NT1M.

††MSE's were estimated from the covariance analysis using NT1M as the covariate.

†††These MSE's are average MSE in the covariance analysis based upon Finney's approximation (1946).

Table 4.47 Comparison of analysis of covariance models using NT1M and stand as the covariates with respect to reduction in MSE and relative efficiency for the variable Y_2 in DATA OIL PALM

Covariates in model	Reduction in MSE (%) for		Relative efficiency (%) for	
	Overall F-test	Pairwise treatment comparison	Overall F-test	Pairwise treatment comparison
None vs stand	63.05	-3.30	270.61	96.81
None vs stand after NT1M†				
per-plant data	68.31	11.41	315.53	112.88
per-plot data	69.07	13.53	323.26	115.65
None vs NT1M and stand	69.12	6.62	323.78	107.09
NT1M vs stand after NT1M				
per-plant data	68.81	12.81	320.60	114.69
per-plot data	69.20	13.90	324.65	116.14
NT1M vs NT1M and stand	69.25	25.17	325.19	133.64
Stand vs stand after NT1M				
per-plant data	14.24	14.24	116.60	116.60
per-plot data	16.29	16.29	119.46	119.46
Stand vs NT1M and stand	16.42	9.60	119.65	110.62

†Stand after NT1M refers to the covariance analysis of plot totals adjusted for NT1M, estimated from per-plant data or per-plot data, using stand as the covariate.

A double covariance analysis can be performed in lieu of the second approach using NT1M and stand as the covariates. This analysis gave a bias of -5.77 ± 3.06 kg of bunches and an estimate of yield per standing plant of 132.81 ± 24.14 kg of bunches. The estimates of yield per standing plant were 133.16 ± 23.48 kg of bunches for the first approach, 132.72 ± 23.20 kg of bunches for the second approach, and 131.67 ± 26.39 kg of bunches for the adjustment for stand only.

For each analysis, one degree of freedom was subtracted from total and error degrees of freedom for each of the covariates defined to measure the missing plant effects and stand differences.

5. SUMMARY

The effects of missing plants upon their adjacent standing plants in field experiments on perennial plants was investigated in this thesis. The readings on the response variables for these standing plants may be altered by missing plants and, therefore, a method to estimate the missing plant effects and to adjust treatment means for resulting biases is needed.

The data from several experiments were analyzed by a per-plant covariance analysis approach using regression techniques. The average amount (positive or negative) and direction of the adjustment on a per-plant basis was estimated by the regression coefficients of the covariates in a particular model.

Two methods are presented for the derivation and designation of the covariates. In the first method, covariates depending upon location were derived; some covariates defined the number and location of the surrounding missing plants to a standing plant. There were unique covariates for vertical, horizontal, and diagonal locations with respect to a standing plant. According to the second method, covariates independent of location of the missing plants with respect to a standing plant were derived; some covariates defined the number and closeness of the surrounding missing plants to a standing plant. In both methods, some covariates were indicators of the presence or absence of missing plants, independent of the number of missing plants. The methodology has been developed for a square planting pattern but similar covariates can be defined for other planting patterns.

Two experiments on loblolly pine (Pinus taeda L.) were used in this study. For these, the response variables were total height of the tree to the nearest 0.1 foot (HT), diameter at breast height measured on the outside bark to the nearest 0.1 inch (DBH), and volume outside bark in cubic feet (VOL). One of these experiments was a heritability study on 28 open-pollinated families of loblolly pine in Decatur County, Georgia (DATA LOBLOLLY PINE 1), and the other was a progeny test on 40 open-pollinated families of loblolly pine in Martin County, North Carolina (DATA LOBLOLLY PINE 2). Data from an oil palm [Elaesis guineensis Jacq.) experiment conducted at Santo Domingo, Ecuador, were also used in this study (DATA OIL PALM). The experiment was designed to compare eight crosses of oil palm in different years (i = 1972, 1973, 1974, and 1975). The response variables in this experiment were as follows: plant yield, measured as the number of bunches per year (B_i), total weight of plant yield per year (Y_i), and mean weight per bunch for each plant in each year (MYB_i). A growth factor experiment on peach (Prunus persica L.) was conducted at Jackson Springs, North Carolina, for several years--1969, 1970, 1971, and 1972 (DATA PEACH). For this, the response variables were growth of the tree as measured by annual increase of basal cross-sectional area of the tree trunk approximately one foot from the soil line (TRAR), yield of fruit by harvested date (TOH), and number of fruit harvested per tree (NOFR).

It was shown clearly that the effects of missing neighboring plants can be measured by the covariates described herein. The covariates independent of location of missing plants were more appropriate for studying the missing plant effects. For almost all response variables, the covariate NT1M defined as the number of missing plants nearest to a standing plant, excluding those missing plants in a diagonal position with respect to the standing plant, was of greatest importance. The exception was for DBH in DATA LOBLOLLY PINE 2 for which the covariate LTM defined as the natural logarithm of (the total number of missing plants surrounding a standing + 1) was better.

The results using LTM as a covariate for DBH in DATA LOBLOLLY PINE 2 were as follows:

1. The estimated overall per-plant bias was -0.08 $\pm$ 0.04 inches.

2. The largest estimated total bias for a treatment effect was -5.99 inches.

3. Missing plants occurred at 15.19% of the plant positions in the entire experiment. An estimate of the total bias caused by all of these missing plants was -150.05 inches.

4. The average bias for the entire experiment expressed as a percentage of the adjusted overall experiment mean was -1.30%.

5. The percent reduction in experimental error variance was 1.98%.

It is important to indicate that the effects of NT1M as a covariate for DBH in DATA LOBLOLLY PINE 2 were marginally significant with (Prob > F) = 0.0602 whereas (Prob > F) = 0.0445 for LTM.

The results using NT1M in the covariance model were as follows:

1. The estimated overall per-plant biases were: 0.20 $\pm$ 0.05 inches for DBH and 0.36 $\pm$ 0.09 cubic feet for VOL in DATA LOBLOLLY PINE 1 (HT was nonsignficant); -0.30 $\pm$ 0.10 feet for HT and -0.03 $\pm$ 0.01 cubic feet for VOL in DATA LOBLOLLY PINE 2; -0.44 $\pm$ 0.19 bunches for B_1, -7.49 $\pm$ 3.11 kg of bunches for Y_2, and -0.59 $\pm$ 0.30 kg per bunch in DATA OIL PALM; and 70.08 $\pm$ 33.16 square inches for TRAR (1971), 8.46

$\pm$ 3.34 pounds of fruit for TOH (1971), and 6.22 $\pm$ 2.72 pounds of fruit for TOH (1972).

2. The largest estimated total biases for a treatment effect were: 11.73 inches for DBH and 21.06 cubic feet for VOL in DATA LOBLOLLY PINE 1; -18.03 feet for HT and -1.65 cubic feet for VOL in DATA LOBLOLLY PINE 2; -883.92 kg of bunches for Y_2, -69.72 kg per bunch for MYB_2, and -52.12 bunches for B_1 in DATA OIL PALM; and 5676.31 square inches for TRAR (1970), 12738.66 square inches for TRAR (1971), 922.20 pounds of fruit for TOH (1971), and 745.88 pounds of fruit for TOH (1972) in DATA PEACH.

3. Missing plants occurred at 19.19%, 15.19%, and 13.37% of the plant positions in each of the entire experiments for DATA LOBLOLLY PINE 1, DATA LOBLOLLY PINE 2, and DATA OIL PALM, respectively. Estimates of the total bias caused by all of these missing plants were 247.96 inches for DBH and 445.18 cubic feet for VOL in DATA LOBLOLLY PINE 1; -473.00 feet for HT and -43.23 cubic feet for VOL in DATA LOBLOLLY PINE 2; and -2419.59 kg of bunches for Y_2, -190.85 kg per bunch for MYB_2, and -142.66 bunches for B_1 in DATA OIL PALM. For DATA PEACH, missing plants occurred at 22.66%, 27.60%, and 47.14% of the plant positions in the entire experiment in 1970, 1971, and 1972, respectively. Estimates of the total bias caused by all of these missing plants were 11352.62 square inches for TRAR (1970), 24191.76 square inches for TRAR (1971), 1751.33 pounds of fruit for TOH (1971), and 1367.45 pounds of fruit for TOH (1972).

4. The average biases for each of the entire experiments expressed as a percentage of the adjusted overall experiment means were: 2.31% for DBH and 5.19% for VOL in DATA LOBLOLLY PINE 1; -0.79% for HT

and -2.15% for VOL in DATA LOBLOLLY PINE 2; -4.11% for Y_2, -3.01% for MYB_2, and -5.28% for B_1 in DATA OIL PALM; and 3.24% for TRAR (1970), 5.84% for TRAR (1971), 5.41% for TOH (1971), and 10.41% for TOH (1972) in DATA PEACH.

5. The percent reductions in experimental error variance were: 14.11% for DBH and 11.26% for VOL in DATA LOBLOLLY PINE 1: 5.15% for HT and 3.22 for VOL in DATA LOBLOLLY PINE 2; 4.75% for Y_2, 7.45% for MYB_2, and 5.40% for B_1 for DATA OIL PALM. For DATA PEACH, there were six error terms involved in the analysis of the data; the MSE's for these error terms varied among and within year; the extreme reductions in MSE were -31.82% for E(d) and 26.42% for E(e).

To illustrate the subsequent analysis of the data after adjusting for the missing plant effects, a covariance analysis of plot totals adjusted for the missing plant effects using stand differences as a covariate was carried out for the variable Y_2 in DATA OIL PALM. This covariance analysis showed the effectiveness of adjusting for both the missing plant effects and stand. In this case, there was a larger reduction in MSE (68.99%) and relative efficiency (115.38%) in relation to the covariance analysis using only stand as a covariate. The latter analysis had a 63.05% reduction in MSE and a 96.81% relative efficiency. An approximation to the estimation of the per-plant missing plant effects may be obtained by a per-plant covariance analysis where the covariate is defined according to the methodology of this thesis. This analysis for the Y_2 variable in DATA OIL PALM gave a bias of -5.35 $\pm$ 5.52 kg of bunches whereas the analysis on a per-plant basis gave a bias of -7.49 $\pm$ 3.11 kg of bunches. If an approximation is needed, a double covariance analysis using the covariates to define the missing plant effects and stand differences can be performed to complete the analysis of the data. For this case, the estimated bias was -5.77 ± 3.06 kg of bunches.

From the results of this research, it was clear that missing plant effects varied with certain agronomical and/or environmental conditions, the response variable, and the covariates used to detect the missing plant effects. Thus, the bias on a per-plant basis increased (positive bias) or decreased (negative bias) the response of standing plants surrounded by missing plants in different amounts. These amounts were more noticeable when the estimated total bias for the treatments and the entire experiment were computed. The covariates reduced the experimental error variance except for the peach experiment. The increase in error for this experiment could be attributed to the many design variables and the six error terms involved in the analysis of DATA PEACH.

The effects of missing plants are expected to be of greatest importance in experiments on crops having plant characteristics such that there will be competitive effects at normal spacings. The competitive effects, although present in the data sets studied, did not appear to be sizable at these spacings which were considered to be normal. One probably could use the techniques described herein to greater advantage in highly competitive situations such as pine nursery seedlings.

When initiating a study on missing plants using the approach described herein, it seems advisable to first use only the covariate NT1M. The results of such an analysis would give some indication of whether the effects of the missing plants on adjacent plants are detectable. Given an indication of an effect, future refinement of the covariates could be carried out. Thus missing plants on diagonal positions and those located at more than one missing plant away from a standing plant may be included as covariates. If a researcher considers location of missing plants important, covariates defined in this thesis that measure the number of missing plants next to a standing plant for each of the vertical, horizontal, and diagonal locations may be used. These covariates can easily be defined to account for the closeness of the missing plants to a standing plant along any location.

In this thesis, the analysis of covariance has been performed on a per-plant basis, where the missing unit is a plant. However, there might be some situations where the missing unit may consist of a certain number of missing contiguous plants that occurred during the experiment. The missing unit could also be measured by the area of the gap. This could be useful for small plant crops of high plant population densities.

The covariates presented in this thesis are independent of time of loss because information about the dates on which plants first became missing was not recorded. One possibility in future research is to weigh a missing plant by the length of time that it has been missing, because missing plants may differ in their effect according to time of loss.

The per-plant covariance analysis presented in this thesis is a general approach for estimating biases caused by missing plants on a per-plant basis. This procedure can be applied in conjunction with practically any experimental design under a variety of agronomical and/or environmental variations presented during the conduct of the experiment. The intensity of these effects will be reflected in the amount and direction of the estimated regression coefficients for the covariates.

The per-plant covariance is intended to replace the approximate formulas for estimating effects of missing plants on adjacent plants. The advantage here is that adjusting is done using a factor (regression coefficient) estimated from that particular set of data rather than a general formula applying to all fields of a species.

6. LIST OF REFERENCES

Abeywardena, V. 1964. Statistical control of variability in coconut experiments. Empire J. Expl. Agri. 32:167-174.

Allan, R. E. and J. Wishart. 1930. A method of estimating the yield of a missing plot in field experimental work. J. Agr. Sci. 20:399-406.

Blake, G. M. 1959. A study to determine optimum plot size for progeny testing of red pine. M. S. thesis, University of Minnesota, Minneapolis.

Brewbaker, H. E. and F. R. Immer. 1931. Variations in stand as sources of experimental error in yield tests with corn. J. Amer. Soc. Agron. 23:469-480.

Bush, R. H. and Y. Ergun. 1973. Yield adjustment for missing units in spring wheat. Crop Sci. 13:126-128.

Chapas, L. C. 1961. Plot size and reduction of variability in oil-palm experiments. Empire J. Expl. Agri. 29:212-224.

Cochran, W. G. 1957. Analysis of covariance: Its nature and uses. Biometrics 13:261-281.

Cochran, W. G. and G. M. Cox. 1966. Experimental Designs. John Wiley and Sons, Inc., New York.

Coons, I. 1957. The analysis of covariance as a missing plot technique. Biometrics 13:387-405.

Crews, J. W. and G. L. Jones. 1962. Procedure for adjusting yields on the basis of stand in flue-cured tobacco experiments. Tobacco Sci. 6:114-118.

Federer, W. T. 1957. Variance and covariance analyses for unbalanced classifications. Biometrics 13:333-362.

Finney, D. J. 1946. Standard errors of yield adjusted for regression on an independent measurement. Biometric Bull. 2:53-55.

Fisher, R. A. 1932. Statistical Methods for Research Workers. 4th Ed. Oliver and Boyd, Edinburgh, Great Britain.

Fitch, C. L. and E. R. Bennett. 1910. The potato industry in Colorado. Colo. Agr. Expt. Sta. Bull. 175, pp. 65-68.

Giesbrecht, J. 1961. Effect of incomplete hills and compensating treatments in comparative corn yield trials. Canad. J. Plant Sci. 41: 91-96.

Gomez, K. A. and S. K. De Datta. 1972. Missing hills in rice experimental plots. Agron. J. 64:163-164.

Gupton, C. L. and L. E. Archer. 1973. Procedure for adjusting the yield of plots of burley tobacco (Nicotiana tabacum L.) for differential stands. Agron. J. 65:101-104.

Haines, W. B. and B. Benzian. 1956. Some manuring experiments on oil-palm in Africa. Empire J. Expl. Agri. 24:137-160.

Hindi, L. H. A. 1962. The effect of skips on grain yield of the adjacent hills in trials with maize hybrids under different conditions. Euphytica 11:327-356.

Hoyle, M. H. 1971. Spoilt data--An introduction and bibliography. J. R. Statist. Soc. A 134:429-439.

Jones, G. L. and W. K. Collins. 1959. Measure crop performance: Tobacco. Dept. Field Crops, Res. Rept. 19, N. C. State University, Raleigh.

Kiesselbach, T. A. 1918. Studies concerning the elimination of experimental error in comparative crop tests. Nebr. Agri. Expt. Sta. Res. Bull. 13.

Kiesselbach, T. A. 1923. Competition as a source of error in comparative corn yields. J. Amer. Soc. Agron. 15:199-215.

Kowal, J. M. L. 1959. The effect of spacing on the environment and performance of cacao under Nigerian conditions. I. Agronomy. Empire J. Expl. Agri. 27:27-34.

Mahoney, C. H. and W. D. Baten. 1939. The use of the analysis of covariance and its limitations in the adjustment of yields based upon stand irregularities. J. Agri. Res. 58:317-328.

Moursi, M. A. 1956. A note on the relationship of gaps in stand and yield of dasheen. Empire J. Expl. Agri. 24:245-246.

Myers, C. H. and F. R. Perry. 1923. Analysis and interpretation of data obtained in comparative tests of potatoes. J. Amer. Soc. Agron. 15:239-253.

Pearce, S. C. 1953. Field experimentation with fruit trees and other perennial plants. Commonw. Bur. Hort. and Plantation Crops, Tech. Comm. 23.

Pearce, S. C. 1955. Some considerations in deciding plot size in field trials with trees and bushes. J. Ind. Soc. Agri. Stat. 7:23-26.

Pope, O. A. 1947. Effect of skips, or missing row segments, on yield of seed cotton in field experiments. J. Agri. Res. 74:1-13.

Scheffé, H. 1959. The Analysis of Variance. John Wiley and Sons, Inc., New York.

Searle, S. R. 1971. Linear Models. John Wiley and Sons, Inc., New York.

Service, J. 1972. A User's Guide to the Statistical Analysis System. Dept. of Stat., N. C. State University, Raleigh.

Smith, H. F. 1957. Interpretation of adjusted treatment means and regressions in analysis of covariance. Biometrics 13:282-308.

Snedecor, G. W. and W. G. Cochran. 1968. Statistical Methods. The Iowa State University Press, Ames.

Steel, R. G. D. and W. T. Federer. 1955. Yield-stand analyses. J. Ind. Soc. Agri. Stat. 7:27-45.

Steel, R. G. D. and J. H. Torrie. 1960. Principles and Procedures of Statistics. McGraw-Hill Book Co., Inc., New York.

Stewart, F. C. 1919. Missing hills in potato fields: Their effect upon the yield. N. Y. Agr. Expt. Sta. (Geneva) Bull. 459.

Stonecypher, R. W., B. J. Zobel and R. Blair. 1973. Inheritance patterns of loblolly pines from a nonselected natural population. N. C. Agr. Expt. Sta. Tech. Bull. 220.

Teigen, J. B. and J. J. Vorst. 1975. Soybean response to stand reduction and defoliation. Agron. J. 67:813-816.

Wilkinson, G. N. 1960. Comparison of missing value procedures. Aust. J. Stat. 2:53-65.

Woessner, R. A. 1965. Growth, form, and disease resistance in four-year-old control- and five-year-old open-pollinated progeny of loblolly pine selected for use in seed orchards. School of Forest Resources, Tech. Rept. 28, N. C. State University, Raleigh.

Wright, J. W. and F. D. Freeland. 1960. Plot size and experimental efficiency in forest genetic research. Mich. Agr. Expt. Sta. Tech. Bull. 280.

Yates, F. 1933. The analysis of replicated experiments when the field results are incomplete. Empire J. Expl. Agri. 1:129-142.