Spatial Patterns of Tree Height Variations in a Black Spruce Farm-Field Progeny Test and Neighbors-adjusted Estimations of Genetic Parameters

By D. JOYCE[1] [4], R. FORD[2] and Y. B. FU[3]

(Received 14th September 2001)

Summary

A farm-field test of black spruce progeny with random non-contiguous single tree plots was analyzed with spatial techniques and nearest-neighbors adjustments to evaluate the effectiveness of used blocking and neighbor adjustments in controlling the site heterogeneity for estimations of genetic parameters. Spatial analyses of variation in tree height at ages 3-10 showed that row and column displayed different shapes of gradients (roughly V and N shapes, respectively) after age 6 and these gradients together explained 9.7% of the site variation at age 10. Also, the patchy structure found followed an exponential covariance model with an estimated range of 12 plots (i.e., 7–11 m) across and accounted for 19.8% of the site variation at age 10. The applied blocking effectively removed 22.7% of the removable site variation at age 10, but posterior data adjustments with 4, 8 and 12 nearest neighbors removed 57.6, 77.6 and 83.1% of it, respectively. The neighbor adjustments also displayed considerable impacts on estimates of genetic parameters associated with family rankings and genetic gains of family, individual and early selections. The standard errors of a difference between family mean estimates and the estimated variances of all the family means were reduced up to 10.2 and 19.3%, respectively. The estimates of family and individual heritabilities and the estimated efficiencies of early selection were increased up to 6.7, 16.4 and 16.6%, respectively. These findings are encouraging for future uses of spatial techniques and posterior neighbor adjustments in analyses of forest genetic trials.

Key words: Picea mariana, spatial analysis, nearest-neighbors analysis, heritability, age-age correlation, early selection.

Introduction

Black spruce (*Picea marianna* [MILL.] B.S.P) is one of the most important reforestation species with an annual planting of 65 million trees in northern Ontario, Canada. Genetic improvement of black spruce in Ontario was started in the early 1980's with the selection of approximately 8,000 plus trees. These plus trees were later allocated into 27 breeding programs. Subsequent measurements of progeny tests have been used for incremental roguing of the associated seedling seed orchards. In recent years, both industry and the provincial government have initiated the second generation breeding programs to mitigate some shortfalls in wood supply resulting from increases in amount of lands set aside as parks and protected areas.

In 1990, two field tests of black spruce progeny for one breeding program in northeastern Ontario were written off due to poor survival and unexpectedly high heterogeneity within the test sites. A farm-field test was thus established in 1992 with the goal to generate some genetic information for roguing the associated first generation seedling seed orchard. Farm-field testing can generate high quality genetic information at young ages through extensive site preparations, close spacing and nearly complete weed control. Over the last decade, the use of farm-field testing in Canada has substantially increased (CARLSON, 1990), but few studies on the effectiveness of farm-field tests in operational programs have been made (WOODS et al., 1995). Information is needed on the control of site heterogeneity, the quality of genetic parameter estimates and the relative efficiency of early selections for growth potential over those in regular field tests to support decisions regarding advanced generation breeding programs.

In this study, we performed a detailed analysis of the Aidie Creek farm-field test of black spruce progeny with the hope to understand (1) how homogeneous the test site had been through extensive site management; (2) how effective the applied blocking had been in controlling the site variation; and (3) whether an application of posterior data adjustments can enhance the control of the site variation to achieve its test goal. Specifically, spatial techniques were first applied to describe the site variation and to estimate the patch size and the experimentally removable site variation (FU et al., 1999). These estimates allowed us to derive the maximum of nearest-neighbors to be applied in neighbor adjustments and evaluate the effectiveness of the applied blocking and neighbor adjustments in controlling the site variation, respectively. Then the test data were subject to the derived nearest-neighbors adjustments before estimations of various genetic parameters were made to assess the effects of the neighbor adjustments.

Materials and Methods

Test materials and data collection

The farm-field test used in this study was established between June 3 and 9, 1992, at the Swastika Tree Nursery (Lat. 48°07' and long. 80°06'), Swastika, Ontario, a representative of the severe end of the climatic variation in the associated breeding zone of black spruce. The test consisted of 397 open-pollinated black spruce families each having 30 seedlings. The container stock was seeded between March 4 and 8, 1991, grown in Leach tubes, and over-wintered in frozen storage at –4°C. The applied field design followed random non-contiguous plots in interlocking field layouts (LIBBY and COCKERHAM, 1980), with 10 complete blocks each having three interlocking replications of 397 single tree plots. The spacing of the trees was 60 cm between trees within rows and 90 cm between rows. No surround trees were planted. As an active nursery field, the site had been well tilled and was flat and well drained. The soil type was a sandy-loam texture that seemed to be consistent throughout the test site. Clean cultivation was made with chemical and manual methods for the first two years of planting.

---

[1] Ministry of Natural Resources, 70 Foster Drive, Suite 400, Sault Ste. Marie, Ontario P6A 6V5, Canada

[2] Northeast Seed Management Association, Box 730, Englehart, Ontario P0J 1H0, Canada

[3] Plant Gene Resources of Canada, AAFC Saskatoon Research Centre, 107 Science Place, Saskatoon, Saskatchewan S7N 0X2, Canada

[4] Corresponding author:
Dr. DENNIS JOYCE, Ministry of Natural Resources, 70 Foster Drive, Suite 400, Sault Ste. Marie, Ontario P6A 6V5,
Canada. Phone: 705-945-5894, Fax: 705-945-6667, Email: Dennis.Joyce@mnr.gov.on.ca

Field assessments were made annually from 1993 to 2000 for tree height and growth condition. The trees with severe damages and the four outliers (i.e., those observations greater than three standard errors) detected from the normality tests of trait value distribution were treated as missing in the analyses below.

Spatial analyses of residuals

To evaluate the spatial patterns of site variation in the farm field test, we analyzed the residuals of the tree heights measured over the eight ages. For the tree heights of a given age, family means were calculated and deducted from each observation of the family member, which was done with SAS PROC GLM (SAS® Institute Inc., 1996). The resulting residuals, however, should still retain ¾ of the genetic variation, as within-family genetic variation can not be effectively removed from such an open-pollinated progeny test data. These residuals could also be confounded with block and replication effects.

Residuals of a given age were first analyzed with a median-polishing technique to (*i*) obtain the medians of the residuals for each row and column for the analyses of trend (or gradients) and (*ii*) generate detrended residuals by removing these medians from rows and columns for the analyses of the small-scale spatial structure (FU et al., 1999). Note that medians rather than means were used here because the former is more robust than the latter to unequal numbers of observations available among rows or columns. Also note that the median-polishing only by rows and columns, not in the other orientations, may not capture all of the large-scale deterministic structure. In this study, median and detrended data were first generated with a program written in SAS IML (SAS® Institute Inc., 1996). Then the medians over rows and columns were plotted to evaluate the gradients. The plotting was made for the residual medians of all the eight ages to assess their temporal stability.

Detrended residuals of a given age were analyzed to obtain nugget, sill and range to characterize patch variations of the test site using a variogram technique (MATHERON, 1963). This technique and its application to forest genetic test data were discussed in detail by FU et al. (1999). The main idea of the technique is first to obtain experimental variances as a function of distances between plots and then to fit these experimental variances with a theoretical model, so that nugget, sill, and range can be estimated. Specifically, the lag distance where the modeled variance approaches an asymptotic maximum (defined as a sill) estimates the range across which data are spatially correlated. As the lag distance approaches zero, the modeled variance also approaches a finite value, called the nugget. Note that the sill is equal to the sum of the nugget and the patch variance. In this study, SAS PROC VARIOGRAM (SAS® Institute Inc., 1996) was used to obtain the isotropic variogram over a maximum of 60 lag distances. Attempts to fit the experimental variogram into various theoretical spatial models (such as spherical, exponential, Gaussian) were made with SAS PROC NLIN (SAS® Institute Inc., 1996). It was found that the following exponential spatial model mostly explained the experimental variogram [$2\gamma(h)$]:

$$2\gamma(h) = n + p(1 - e^{-h/a}), \tag{1}$$

where n is the nugget, p the patch variance, h the lag distance, and a the range. Fitting the exponential spatial model to the experimental variogram gave the estimates of nugget, sill, and range. To illustrate the fitting, the experimental and fitted variograms were plotted, along with the estimates of patch variance, nugget, and range. The plotting was repeated for the variograms of all the eight ages to assess their temporal stability.

Neighbor adjustments

In this study, we adjusted the height measurement of a subject tree from those of 4, 8, and 12 nearest neighbor trees (N4, N8, and N12, respectively, for brevity) as below:

N4	N8	N12
		x
x	x x x	x x x
x s X	x s x	x x s x x
x	x x x	x x x
		x

where s is the subject tree and x stands for a selected nearest neighbor tree. These chosen neighborhood sizes were largely based on the estimates of patchy size from the spatial analysis and expected to be smaller than the average patchy size detected in this study (see below). For each neighborhood size, the height residual (at a given age) of the subject tree was adjusted by the average height residual for all the selected nearest neighbor trees. A down adjustment was made for the subject tree if the average height residual was positive (i.e., an above-average neighborhood) and an up adjustment when the average height residual was negative. Note that the number of the selected nearest neighbor trees for a subject tree can be equal or less than the specified neighborhood size because some selected trees may have missing values or the subject trees may be located on the marginal rows or columns. The adjusted residual was then added with its corresponding family mean. This procedure was repeated for each tree in the test. It was done with a computer program written in SAS IML (SAS® Institute Inc., 1996) and one data file was generated for each neighborhood size for later analyses.

Estimation of genetic parameters

To assess the effects of the neighbor adjustments used, we considered (1) the total phenotypic variance, (2) the standard error of a difference between family mean estimates, (3) the error variance of family mean estimates, (4) the family variance, (5) the family heritability, (6) individual heritability, and (7) age-age correlation. These criteria are sensitive to the posterior removing of site variation, the ranking of families, the genetic gains of family and individual selections, and the efficiency of early selection. Specifically, the first criterion can be considered to be an indicator of its effectiveness in removing the gradient and patchy variations of the test site. The 2nd and 3rd would measure the precision of both estimating family means and ranking the families. The 4th would measure directly the changes of additive genetic variance. The 5th and 6th would affect the calculations of genetic gains for family and individual selections. The 7th was directly linked to the efficiency of early selection.

For the tree heights of each age, the adjusted and unadjusted data were first subject to analysis of variance components with SAS PROC VARCOMP (SAS® Institute Inc., 1996) and estimation of family means with SAS PROC MIXED (SAS® Institute Inc., 1996), in which REML options were used. The following linear model was used:

$$Y_{ikl} = U + B_i + F_k + e_{ikl}, \tag{2}$$

where Y_{ikl} is the observed value of the l^{th} tree in the k^{th} family growing in the i^{th} block, U is the overall mean, B_i is the effect of the i^{th} block (i=1..10), F_k is the effect of the k^{th} family (k=1..397), and e_{ikl} is the random error due to the l^{th} tree in the k^{th} family growing in the i^{th} block. The family-block interaction was omitted in the model, because of insufficient computing RAM memory; specifically, the inclusion of the interaction term required up to 500MB RAM memory, depending on specific analysis. Attempts were made to analyze the significance of the interaction term with SAS PROC GLM (SAS® Institute Inc., 1996) by removing 97 families randomly from the data without neighbor adjustment. The GLM analyses of the selected data detected no significant block-family interaction at the 0.05 level, eight out of 10 trials. The two significant trials showed only marginal significance (at the 0.039 and 0.048 levels). This indicated that the family-block interaction in this test was probably marginally significant. Thus the exclusion of the interaction term could result in some bias in estimations of various genetic parameters, but such bias may not have significant impacts on the comparative assessments of effectiveness. For the estimation of family means, the block effect in the model (2) was treated as random. Analysis of variance components was performed with both random and mixed models; in the random model, both block and family were treated as random effects while in the mixed model, the block effect was treated as fixed. The analysis with the random model provided estimates of block variance to evaluate the effectiveness of the used blocking in controlling the site variation.

From the analysis of variance components with the mixed model for the tree heights of each age, the total phenotypic variance and family variance were obtained and the family and individual heritabilities (h_f^2 and h^2) were estimated as follows:

$$h_f^2 = \sigma_f^2 / (\sigma_f^2 + \sigma_e^2 / n_f) \text{ and } h^2 = 4\sigma_f^2 / (\sigma_f^2 + \sigma_e^2) \tag{3}$$

where σ_f^2 is the family variance, σ_e^2 is the error variance, and n_f is the average family size. From the estimates of 397 family means, the error variance of family mean estimates and the standard error of a difference between family mean estimates were calculated. These six assessment criteria were obtained for the tree heights of each age without and with the neighbor adjustments. The ratios of the estimates for these criteria from the adjusted data over those without adjustment were also calculated for ease in comparison.

Additive genetic correlation of the tree heights at two different ages ($R_{y,Y}$) was estimated from the analyses of family variance for the tree heights of each age and their cross products over the eight ages as:

$$R_{y,Y} = \sigma_{yY}^2 / \sqrt{(\sigma_y^2 \sigma_Y^2)}, \tag{4}$$

where σ_{yY}^2 is the family covariance component between younger and older traits y and Y, the other two variances are the family variances for younger and older traits y and Y, respectively. A linear regression analysis was made of the genetic age-age correlations obtained for the eight ages over the natural logarithm of the age ratio (younger age/older age) following LAMBETH (1980). The analysis generated a predicted $R_{y,Y}$ used to estimate the efficiency of early selection (E) as:

$$E = \frac{h_y^2 R'_{y,Y} T_Y}{h_Y^2 T_y}, \tag{5}$$

where h^2's are the individual heritabilities and T's are the generation times for younger and older traits y and Y, respectively. The relative efficiency of early selection with neighbor adjustment over that without adjustment (RE) was estimated as:

$$RE = \frac{h_y^2 R'_{y,Y}(adjusted)}{h_y^2 R'_{y,Y}(unadjusted)}. \tag{6}$$

This relative efficiency takes into account the estimates of individual heritability with neighbor adjustment. To illustrate the impact of the neighbor adjustments on RE, we considered a juvenile selection at ages 3–10 and compared it to a relatively mature selection at age 20 (i.e., roughly half the rotation age of black spruce). We also assumed four additional years for breeding time and the predicted individual heritability at age 20 as 0.20. Note that invalidity of this assumption should not have much effect on the comparison of selection efficiencies.

Results

Spatial patterns of tree height variations

Evaluations of the residual medians of tree height at ages of 3 to 10 over 60 columns and 200 rows showed that the gradients in both column and row directions were not linear; they appeared to be an N-shape in column direction and a V-shape in row direction. These patterns became obvious when the trees were six years old and did not change much at the older ages (*figure 1*). The patterns of experimental variances as a function of lag (plot) distances and the fittings of the detrended residuals with an exponential covariance model are displayed in *figure 1* for the trees at ages 3, 6 and 10. It is clear that the experimental variograms were fitted well with the used exponential covariance model for trees with age 6 and older. As expected, the estimates of patch variation increased over the ages, so did nugget (or plot variance). The estimates of range greatly fluctuated from 6.3 to 13.4 over the first five years of tree growth and did not change much at the older ages, being around 12 plots (with standard errors ranging from 0.66 to 0.72). This means that the patchy size would be 7–11 meters across (noting that the spacing of the trees in rows and columns was not equal as described above). The proportions of the total phenotypic variance explained by gradient, patchiness, and plot were given in *Table 1.* The gradients in both row and column explained 9.7% of the total phenotypic variance and the patchiness accounted for 19.8%, as revealed for H10 (i.e., the tree height of age 10). Combining both the gradient and patchiness variances gave 29.5% of the total phenotypic variance that could be removable, in principle, either by a prior effective field layout (or various blocking) or posterior data analysis (such as neighbor adjustment). Such removable site variation was compatible with those reported from the Douglas-fir progeny trials at southern coastal British Columbia (FU et al. 1999).

Effectiveness of blocking and neighbor adjustment

The proportions of estimated block variance over the total phenotypic variance were given in *Table 1.* Clearly they fluctuated from 0.6 to 10.7% over the eight ages. At age 10, 6.7% of the total phenotypic variation was explained by block. This means that the blocking applied in this test effectively removed 22.7% (=6.7/29.5 in percentage) of the removable site variation. The effectiveness of the three neighbor adjustments in removing the site variation was also given in *Table 1.* For example, the three neighbor adjustments (N4, N8 and N12) removed 17.0, 22.9 and 24.5% of the total phenotypic variance, respectively, as revealed for H10. These accounted for 57.6, 77.6 and 83.1% of the removable site variation, respectively, for

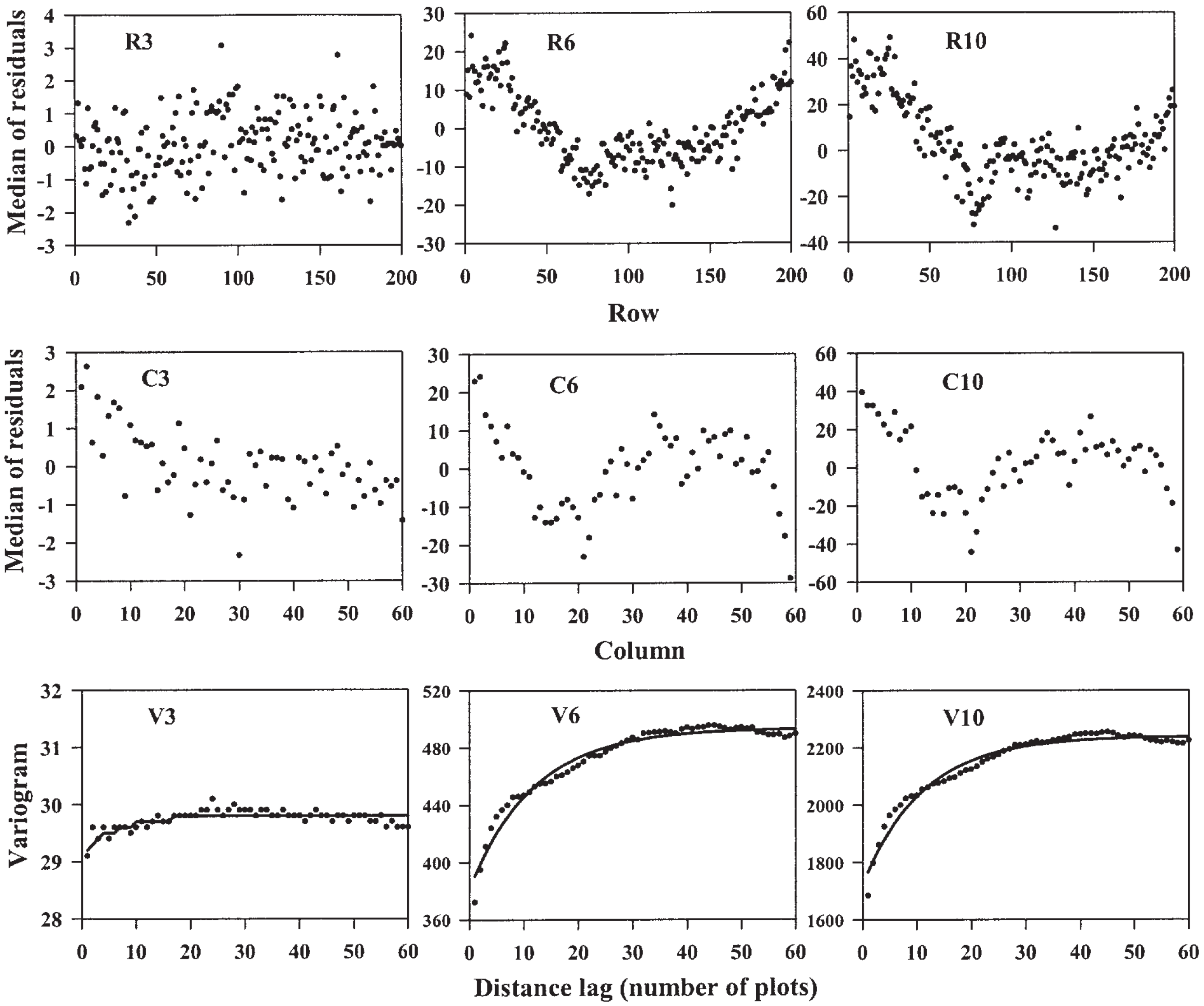

Figure 1. – Spatial variation patterns of tree heights at ages 3, 6, and 10 in the Aidie Creek farm-field test. R3, R6 and R10 show variation patterns of residual medians in row direction for the three ages, respectively C3, C6 and C10 display variation patterns of residual medians in column direction for the three ages, respectively. V3, V6 and V10 reveal experimental variograms and their exponential covariance model fittings of the detrended residuals for the three ages, respectively. The dots over lag distance in each variogram represent the experimental variogram and the fitted line is the modeled variogram. Note that the scales for Y-axis differ for the different ages.

Table 1. – Average height (Ah), total phenotypic variance (P0), the percentage of the total phenotypic variance explained by plot (Pt0), gradient (G0), patch (Pa0), and block (Bk0), without neighbor adjustment, and the percentage reduction of the total phenotypic variance under the three neighbor adjustments (R_N4, R_N8, R_N12).

Age	Ah*	P0	Pt0	G0	Pa0	G0+Pa0	Bk0	R_N4	R_N8	R_N12
3	44.7	31.2	93.1	0.4	2.1	2.5	0.6	-14.7	-6.4	-2.8
4	56.1	47.1	82.9	1.3	4.4	5.7	0.8	-7.1	3.8	6.5
5	72.4	146.1	73.0	14.0	8.2	22.3	2.6	2.9	9.8	12.8
6	120.9	616.1	61.9	14.7	18.2	32.9	9.5	23.3	28.2	29.9
7	154.9	1043.6	60.1	14.4	20.9	35.3	10.2	25.4	30.9	32.3
8	182.2	1652.6	58.9	11.1	24.9	36.0	10.7	21.2	29.6	31.8
9	227.0	2059.8	59.5	10.3	25.2	35.5	8.8	13.1	24.8	27.8
10	262.0	2617.8	65.8	9.7	19.8	29.5	6.7	17.0	22.9	24.5

* Measurement unit for height was cm.

the three neighbor adjustments. However, such effectiveness varied greatly over the other ages.

Effects on family mean estimates

The effectiveness of the neighbor adjustments in removing the site variation was clearly reflected in the reduction of both the standard errors of a difference between family mean estimates and the variances of all the family mean estimates, as given in *Table 2.* For example with H10, the reductions of the standard errors of a difference between family means were 7.6, 8.4 and 10.2%, respectively, for the three adjustments (N4, N8 and N12) and their corresponding reductions of the variances of all the family means were 14.3, 16.0 and 19.3%. Clearly, the neighbor adjustments increased the precision of estimating family means and consequently the precision of ranking families, more with more neighbors adjusted. However, the increases for the precision of estimating family means fluctuated widely with respect to the tree ages (*Table 2*).

Table 2. – The standard error of a difference (SED) between family mean estimates and the variance of all the family mean estimates (VFM) for tree heights of each age without neighbor adjustment (N0) and the percentage reduction of their neighbors-adjusted estimates over those without (R_N4, R_N8, R_N12).

Age	SED				VFM			
	N0	R_N4	R_N8	R_N12	N0	R_N4	R_N8	R_N12
3	4.1	0.1	1.1	1.4	8.6	0.7	2.2	3.0
4	3.9	2.0	3.6	4.3	7.7	4.6	7.3	8.5
5	4.8	1.5	4.3	6.1	11.8	2.4	7.8	11.2
6	10.2	11.3	12.0	13.2	52.7	20.8	22.3	24.1
7	13.0	13.0	13.6	14.8	85.3	23.6	24.9	27.0
8	17.6	11.9	13.7	15.4	155.5	21.5	25.0	27.9
9	20.1	3.9	8.4	10.2	204.0	7.0	15.8	19.2
10	20.7	7.6	8.4	10.2	216.4	14.3	16.0	19.3

Effects on heritability estimates

Estimates of family variance were increased over the ages, but they were slightly reduced with more neighbors adjusted (*Table 3*). Estimates of family heritability without neighbor adjustment ranged from 0.50 to 0.60 (excluding those for the ages of 3 and 4) and were generally increased with more neighbors adjusted. For example with H10, the estimates of family heritability increased from 0.577 (without neighbor adjustment) to 0.591, 0.616 and 0.608 with adjustments of N4, N8 and N12, respectively. Estimates of individual heritability without neighbor adjustment were about 0.20 (excluding those for the ages of 3 and 4), but they generally increased markedly with more neighbors adjusted. At age 10, for example, the estimates of individual heritability increased from 0.191 (without neighbor adjustment) to 0.201, 0.222 and 0.215 for the adjustments of N4, N8 and N12, respectively. These accounted for up to 16% increase in estimates of individual heritability. However, the adjustments from 8 to 12 nearest neighbors showed slight reductions in the estimates of both family and individual heritabilities. Note that the high estimates of family and individual heritabilities at the ages 3–4 were due to the presence of unexpectedly high estimates of family variance.

Table 3. – Estimates of family variance (FV), family heritability (FH) and individual heritability (IH) without neighbor adjustment (N0) and the percentage changes of their neighbors-adjusted estimates over those of no adjustment (C_N4, C_N8, C_N12).

Age	FV				FH				IH			
	N0	C_N4	C_N8	C_N12	N0	C_N4	C_N8	C_N12	N0	C_N4	C_N8	C_N12
3	7.75	97	97	96	0.899	98	99	99	0.993	85	91	94
4	6.06	92	92	91	0.789	96	99	99	0.514	86	95	97
5	6.98	99	94	90	0.588	101	102	102	0.191	102	105	104
6	31.90	81	82	80	0.604	102	106	106	0.207	106	114	114
7	48.14	78	80	78	0.565	102	107	106	0.185	104	116	115
8	84.51	78	79	76	0.539	100	106	105	0.205	99	113	111
9	102.25	98	93	90	0.504	106	111	111	0.199	113	124	124
10	124.69	88	90	85	0.577	102	107	105	0.191	106	116	113

Effects on genetic correlation and selection efficiency estimates

The estimated genetic age-age correlations over the eight ages without neighbor adjustment were given in *Table 4.* The linear regression of these correlations over the selection ages was found to be highly significant with a linear coefficient (and its standard error) of 0.385 (0.042). Similarly, the linear regressions from the three neighbor adjustments (N4, N8 and N12) were also highly significant with linear coefficients (and their standard errors) of 0.433 (0.040), 0.444 (0.046) and 0.439 (0.045), respectively. Estimated efficiencies of early selection in black spruce from age 3 to 10, relative to the selection at age 20, and the percentage changes of the estimated efficiencies from the three neighbor adjustments over those of no adjustment are also given in *Table 4.* The estimated efficiencies could be up to 1.82, depending on the age of early selection. Adjustments of neighborhood size up to 8 increased the efficiencies of early selection, up to 24% depending also on the age of early selection. These findings clearly indicate that a neighbor adjustment can have positive impacts on the predicted age-age correlations and the determination of selection age.

Table 4. – Estimated genetic correlations ($R_{y,Y}$) and estimated efficiencies (E) of early selection in black spruce from age 3 to 10, relative to the selection at age 20, and the percentage changes of the estimated efficiencies from the three neighbor adjustments (RE4, RE8, RE12) over those of no adjustment.

Age	$R_{y,Y}$	E	RE4	RE8	RE12
3	0.601	9.852	81.3	86.1	89.3
4	0.653	4.908	83.6	91.7	94.3
5	0.698	1.751	100.1	102.0	101.4
6	0.739	1.822	104.9	112.7	112.9
7	0.775	1.563	104.3	115.5	114.3
8	0.809	1.665	99.3	112.6	111.0
9	0.840	1.556	113.7	124.3	125.3
10	0.868	1.439	107.0	117.7	113.9
Average			105.8	116.6	115.5

Discussion

The analyses presented here have generated several interesting findings from the Aidie Creek farm-field test. First, the farm field test, even with extensive site management, still exhibited substantial site heterogeneity in gradient and patchy structure. Such heterogeneity was stable after age 6. Second, the estimated removable site variation was 30% of the total phenotypic variation at age 10. The applied blocking and the adjustments with 4, 8 and 12 nearest-neighbors effectively removed 23, 58, 78 and 83% of the removable site variation, respectively. Third, the applied neighbor adjustments displayed considerable impacts on the estimations of various genetic parameters. The estimates of family and individual heritabilities and early selection efficiency were increased up to 7%, 16% and 17%, respectively, from the neighbor adjustments. These findings not only provide us insights into how effectively the Aidie Creek farm-field test had performed in the operational breeding programs of black spruce, but also are valuable for the future development of farm-field tests of black spruce and other forest trees.

Our analyses did not consider the asymmetric neighborhood issue from the unequal spacing used in the test layout. It can be reasoned that such asymmetric layout would affect the estimations of patch size and would also add to some uncertainty in estimations of various genetic parameters from the used neighbor adjustments. Further research on the neighborhood correction should provide some insight on the issue. Also, our analyses did not consider block-by-family interaction due to the availability of limited computing resource. Presence of such interaction would certainly affect the estimations of genetic parameters and consequently the comparisons of various neighbor adjustments. Moreover, we applied the estimated patch size to guide the choices of neighbor size in the neighbor adjustments used and did not explore the neighborhood optimization in detail. The results obtained in this study seem to suggest that the 12 nearest-neighbors used should be close to the optimal, but further study on the choice of neighborhood size should be informative for effective uses of neighbor adjustment. Last, pre-subtraction of family means we applied for the neighbor adjustments could bias estimates of various genetic parameters (WILKINSON et al. 1983; MAGNUSSEN 1993; J. A. LOO-DINKINS and J. H. WOODS unpublished). This represents the major limitation of using neighbor adjustments in analyses of genetic trial data (LOO-DINKINS 1992). Thus caution is needed in interpretations of genetic estimates from the adjusted data.

For an effective application of posterior data adjustments, trial data should be first assessed for various spatial covariance models using SAS MIXED and VARIOGRAM procedures (SAS® Institute Inc., 1996). Then the best fitted spatial model should be integrated directly into the analyses of various genetic parameters as in MAGNUSSEN (1990) and APIOLAZA et al. (2000). Such integrating is expected to remove more patchy variation and have less bias (if any) than the approach employed here, thus more effective. However, the integrating approach requires considerable computing resources and is not workable for a regular test data consisted of more than 10,000 records per test (e.g., 400 families x 30 individuals) until the availability of advanced computing resources. This problem also exists for analyses of large trial data when an REML option is required, even without spatial treatments, but it remains less recognized (ADAMS et al. 1994). All of these challenges open a research avenue for posterior data adjustments.

The findings obtained in this study, although largely specific to the Aidie Creek farm-field test, have some general implications for genetic testing of black spruce and other forest trees. First, substantial site heterogeneity could still be found in a farm-field test, even with extensive site management and uniformity seemingly observed across a test site. Caution is needed in the selection of test sites, the proper site preparation and the effective site management (WOODS et al. 1995). Second, the applied blocking could remove a proportion of a site variation, but applications of more effective field designs such as Alpha designs (WILLIAMS and TALBOT, 1996) may help remove more site heterogeneity for higher efficiencies of genetic estimates (FU et al. 1998). Third, a spatial analysis should not be overlooked for any farm-field tests, nor for regular field tests. Such analysis can generate useful information not only for assessing the effectiveness of field layouts in controlling site variations (FU et al. 1999) but also facilitating posterior data adjustments as illustrated in this study. A spatial analysis can be carried out easily with various SAS procedures such as SAS PROC VARIOGRAM (SAS® Institute Inc., 1996). Fourth, the effectiveness of the neighbor adjustments in removing the site variation observed in this study accords well with those previously reported (WRIGHT, 1977; BONGARTEN and DOWD, 1984; THOMPSON and EL-KASSABY, 1988; LOO-DINKINS 1992; ANEKONDA and LIBBY, 1996). Such effectiveness is encouraging for future uses of posterior data adjustment in analyses of forest genetic trials.

Acknowledgements

We gratefully acknowledge Mr. CHRIS ATACH for his assistance in data collection and assessment and the financial support from Forest Genetics Ontario for the research.

References

ADAMS, W. T., WHITE, T. L., HODGE, G. R. and POWELL, G. L.: Genetic parameters for bole volume in longleaf pine: large sample estimates and influences of test characteristics. Silvae Genet. **43**: 357–366 (1994). — ANEKONDA, T. S. and LIBBY, W. J.: Effectiveness of nearest-neighbor data adjustment in a clonal test of redwood. Silvae Genet. **45**: 46–51 (1996). — APIOLAZA, L. A., GILMOUR, A. R. and GARRICK, D. J.: Variance modelling of longitudinal height data from a *Pinus radiata* progeny test. Can. J. For. Res. **30**: 645–654 (2000). — BONGARTEN, B. C. and DOWD, J. F.: Regression and spline methods for removing environmental variance in progeny tests. In: 19th South. For. Tree Impr. Conf. College Station, TX. pp. 312–319 (1984). — CARLSON, M.: Early testing and selection of lodgepole pine in British Columbia. In: Proceed. of Joint Mtg. West. Forest Genet. Assoc. and I.U.F.R.O. Working Parties S2.02–05, 06, 12, and 14. Olympia, Wa. Aug. 20 to 24. Pp. 4.22–4.36 (1990). — FU, Y. B., CLARKE, G. P. Y., NAMKOONG, G. and YANCHUK, A.: Incomplete block designs for genetic testing: statistical efficiencies of estimating family means. Can. J. For. Res. **28**: 977–986 (1998). — FU, Y. B., YANCHUK, A. D. and NAMKOONG, G.: Spatial patterns of tree height variations in a series of Douglas-fir progeny trials: implications for genetic testing. Can. J. For. Res. **29**: 714–723 (1999). — LAMBETH, C. C.: Juvenile-mature correlations in *Pinaceae* and implications for early selection. For. Sci. **26**: 571–580 (1980). — LIBBY, W. J. and COCKERHAM, C. C.: Random non-contiguous plots in interlocking field layouts. Silvae Genet. **29**: 183–190 (1980). — LOO-DINKINS, J. A.: Field test dsign, Chap 4, pp. 96–139. In: Handbook of Quantitative Forest Genetics. L. FINS, S. FRIEDMAN, and J. BROTSCHOL (Eds). Kluwer Acad. Publ. Dordrecht. The Netherlands. pp. 408 (1992). — MAGNUSSEN, S.: Application and comparison of spatial models in analyzing tree-genetics field trials. Can. J. For. Res. **20**: 536–546 (1990). — MAGNUSSEN, S.: Bias in genetic variance estimates due to spatial autocorrelation. Theor. Appl. Genet. **86**: 349–355 (1993). — MATHERON, G.: Principles of geostatistics. Econ. Geol. **58**: 1246–1266 (1963). — SAS Institute Inc.: SAS user's guide, version 6.12 edition. SAS Institute Inc., Cary, N.C. (1996). — THOMPSON, A. J. and EL-KASSABY, Y. A.: Trend surface analysis of a Douglas-fir provenance-progeny test. Can. J. For. Res. **18**: 515–520 (1988). — WILKINSON, G. N., ECHERT, S. R., HANCOCK, T. W. and MAYO, O.: Nearest neighbour (NN) analysis of field experiments. J. R. Statistical Soc. B. **45**: 151–211 (1983). — WILLIAMS, E. R. and TALBOT, M.: ALPHA+. Experimental designs for variety trials. Version 2.3, Design User Manual. CSIRO, Canberra and SASS, Edinburgh (1996). — WOODS, J. H., KOLOTELO, D. and YANCHUK, A. D.: Early selection of coastal Douglas-fir in a farm-field environment. Silvae Genet. **44**: 178–186 (1995). — WRIGHT, J. W.: An analysis method to improve statistical efficiency of a randomized complete block design. Silvae Genet. **27**: 12–14 (1978).