ON IMPLEMENTATION OF INCOMPLETE BLOCK DESIGNS IN FOREST GENETIC FIELD TRIALS

Yong-Bi Fu

Plant Gene Resources of Canada, Saskatoon Research Centre, Agriculture and Agri-Food Canada, 107 Science Place, Saskatoon, Saskatchewan S7N 0X2 Canada; Telephone: 306-956-7642, Fax: 306-956-7246, Email: fuy@agr.ca

Received July 2, 2002; accepted April 26, 2003

ABSTRACT

Forest genetic field trials are a vital part of forest genetic research and will continue to be critical for advances in tree improvement programs, particularly in advanced generations of tree breeding. Implementation of small blocking represents another attempt toward improving the informativeness of future forest genetic trials, by reducing environmental variation to achieve higher accuracy and precision in estimations of breeding values. Such implementation, however, has not been well justified for its effectiveness and benefit. To this end, a review is made on the findings from our investigation on incomplete block designs (ICBs) and the information available from previous related studies with the hope to increase the confidence of tree breeders in applying small blocking. This review focuses on the limitation of randomized complete block designs, effectiveness of ICBs, heterogeneous nature of site variations, and advantages of small blocking. Also some issues and concerns associated with implementation are discussed, including efficiencies with different estimators and testing materials, choice of ICBs, proper blocking on test site, analysis of data from small blocking, and computer programs. Clearly, implementation of small blocking is not only supported with the revealed reductions of site variation, but also shows promise of potential economic benefits, particularly with the recent advances in computer programs to generate efficient field layouts and to analyse trial data.

Keywords: tree breeding, field design, incomplete block design, site variation, simulation

INTRODUCTION

This paper is dedicated to the memory of Dr. Gene Namkoong, 68, who passed away on Sunday March 3, 2002, at North Carolina, USA. Dr. Namkoong, the 1994 Marcus Wallenberg prize recipient, was a distinguished forest geneticist and helped to lay many theoretical and quantitative foundations for modern forest genetics and tree breeding. I was fortunate to be his last postdoctoral fellow at the University of British Columbia from 1996 to 1998 and was greatly inspired by his broad vision of forest genetics and tree breeding to pursue my own research. Application of small blocking in forest genetic field trials is one of those issues Dr. Namkoong was most concerned with, as the heterogeneity of site variation was not well assessed and the limitations of the currently applied blocking were less recognized. As early as the 1970s, he envisioned that small blocking should be applied to improve the informativeness of future forest genetic trials by reducing environmental variation to achieve higher accuracy and precision in the estimating of genetic parameters (NAMKOONG & ROBERDS 1974). However, such applications did not increase over the last 30 years, at least not in North America. This led to the establishment of a research project for me to investigate the statistical efficiencies of incomplete block designs (ICBs) in forest genetic trials with the hope of justifying the use of small blocking in British Columbia (B.C.) tree improvement programs. Collaborating with Dr. Alvin Yanchuk of the B. C. Ministry of Forests, Dr. Peter Clarke at the University of Natal, South Africa, and Dr. Emlyn Williams at CSIRO, Australia, Dr. Namkoong and I managed to demonstrate the presence of enormous site heterogeneity in existing B.C. Douglas-fir progeny trials, the limitation of commonly used, randomized complete block designs (RCBs) in removing site variation, and the possible gain in statistical efficiency and economic benefit from implementing of ICBs.

In this paper, I will review the findings from our investigation and the information available from previous related studies to argue for applications of small blocking and discuss some issues and concerns associated with its implementation of ICBs in forest genetic field trials. It is my hope that this review will increase the confidence of tree breeders in implementing small blocking in forest genetic field trials.

RATIONALE FOR IMPLEMENTATION OF SMALL BLOCKING

The major objective of a forest genetic field trial is to evaluate family and progeny on various test sites in desired environments for selection of superior genotypes for tree breeding. Such evaluation requires implementation of effective field designs to reduce environmental variation as much as possible to achieve high accuracy and precision of estimating genetic values. Efforts have been made over several decades to investigate and develop efficient field designs. Notable efforts include: the analysis of environmental variability (BATCHELOR & REED 1918), the application of single-tree plots (WRIGHT & FREELAND 1960), the lattice trial of *Pinus patula* in Zimbabwe (BURLEY *et al.* 1966), the development of non-contiguous plots (LIBBY & COCKERHAM 1980), the proposal of unbalanced designs (MCCUTCHAN *et al.* 1985), the simulation of field trials (LOO-DINKINS & TAUER 1987), and the implementation of alpha designs (WILLIAMS & MATHESON 1994). While much has been learnt about the efficiencies of various field designs and the practicalities on field layouts, many forest genetic field trials are not as informative as previously thought (MAGNUSSEN 1993a). Estimates of genotype effects are often found to be seriously inflated and have large standard errors in trials with damaged trees, outliers, competition and microsite effects. Even with obvious advantages of single-tree plots in sampling of environmental variations (LIBBY & COCKERHAM 1980), multiple-tree plots are still widely applied in many field trials (LOO-DINKINS 1992). While the efficient alpha designs have been widely implemented in Australia, South Africa, and Asia (WILLIAMS & MATHESON 1994), such implementation does not seem to be appreciated in North America, in spite of some stimulating studies done by Dr. Gene Namkoong and his colleagues to promote small blocking (e.g., see MCCUTCHAN *et al.* 1985, 1989, FRIEDMAN & NAMKOONG 1987). Clearly, it is time to re-visit field designs for increased efficiencies and to explore alternatives.

Limitation of randomized complete block designs

The randomized complete block design, in which each family appears in each block, has up to recently been the most commonly used field design in forest genetic trials (LOO-DINKINS 1992). This design provides some control of site variability by simple blocking and thus is preferred over the completely randomized design, but in practice its ability to account for site variability is typically limited. In most genetic trials with RCB, there are 100–400 families of 4–10 trees planted in each block and 3–8 blocks per site of 2–10 hectares. Such blocks are quite large in size and can not exclude much environmental variability within blocks.

This limitation of RCB can be easily tested using data of existing progeny trials. Interestingly, however, no such tests have been made for forest trials. To understand the effectiveness of the commonly applied RCBs in reduction of site variation, we studied spatial variation patterns of tree heights at ages from 6 to 12 years in a series of B.C. Douglas-fir progeny trials (FU *et al.* 1999b). These trials were established from 1976 to 1986 on 88 test sites that were widely distributed in southern coastal areas, with the goal of evaluating genetic variances and breeding values for B.C. coastal Douglas-fir. They included eight series of 6-parent-tree disconnected half-diallel tests carried out over 10 years (HEAMAN 1978, YANCHUK 1996). Each of these 8 series was conducted on 11 different forest sites, with each of about 150 full-sib families represented by four-tree row plots (with spacing of 3 meters) in four replicates on each site. Crosses were fully randomized within replicates (i.e., diallels were not blocked in replicates). Measurements of tree height and diameter were made twice over the ages of 6–12 years for most of the sites. In this study, data for tree height from the 66 test sites were analyzed with conventional statistics and geostatistical techniques. It was found that the applied RCBs effectively removed on average 5 % of the site variations and effective blocking could remove up to 24 % of the site variation explained by gradients in row and column and mircosite effects (see FU *et al.* 1999b). Also the applied blocking seemed to remove a little more site variation from larger gradients present in row or column directions (see Fig. 1A–B), but such effectiveness was not associated with larger patch sizes (Fig. 1C), nor on the test sites with larger variation (Fig. 1D).

A similar study was made on an Ontario farm-field test of black spruce (*Picea marianna* [Mill.] B.S.P) progeny with 10 complete blocks of non-contiguous single tree plots and the applied blocking was found to remove only 6.7 % of the site variation (JOYCE *et al.* 2002). This finding, along with those from B.C. Douglas-fir trials, clearly provides empirical evidence that the RCBs used were limited in their ability to control site variation and there is still room to improve the effectiveness of field designs for forest genetic trials. Thus investigation and development of smaller blocking to further reduce site variation are justified.

Effectiveness of incomplete block designs

The idea of small blocking such as ICBs developed in

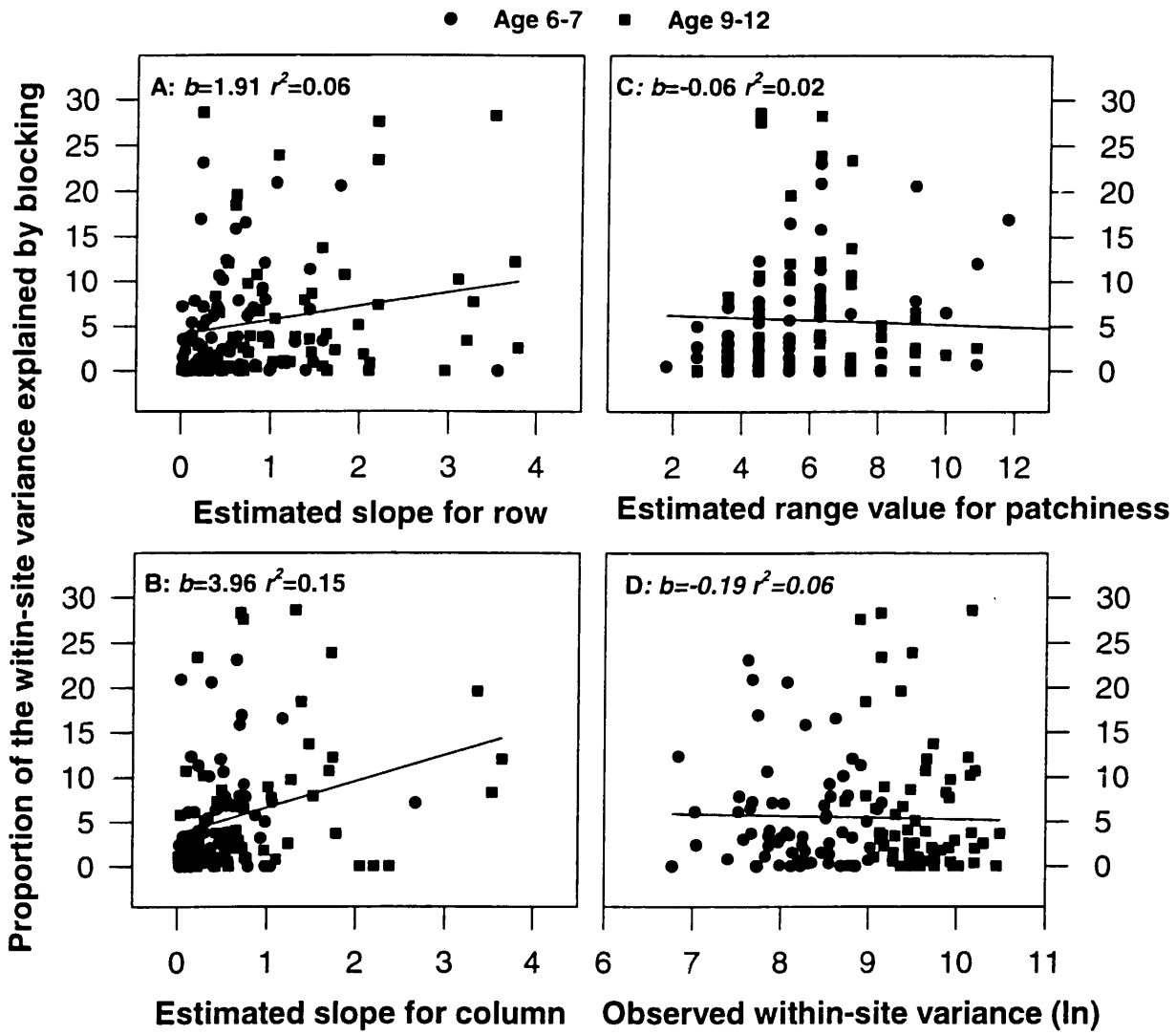

Figure 1. The proportions of within-site variance explained by the applied blocking, in relation to the estimates of slope for both row and column, the estimates of patch size, and the within-site variances of tree height observed twice (at age 6-7 and age 9-12) on the 66 test sites in a series of the Douglas-fir progeny trials (see FU *et al.* 1999b). The linear regressions were made on all the related estimates combined.

the mid-1930's (e.g., see YATES 1936) should be among the alternatives to explore. ICBs subdivide each replicate (presented in one large block under RCBs) into several smaller blocks with each family not necessarily appearing in every block, thus allowing a better control of site variation (COCHRAN & COX 1957, HINKELMANN & KEMPTHORNE 1994, CLARKE *et al.* 1997). Smaller blocks are expected to be less internally heterogeneous than larger blocks, especially on heterogeneous sites, and the larger site variation among blocks is removed from the experimental error so that the contribution of site variance effects to the error of estimating family means can be reduced. Thus, it can be expected, at least statistically, that ICBs are more efficient than RCBs in heterogeneous environments, like those evaluated in forest genetic trials.

Studies of the expected increase in efficiency by small blocking date back to the 1960s when the first progeny trial of *Pinus patula* was conducted in Zimbabwe (BURLEY *et al.* 1966). This trial consisted basically of three replications of ten-tree row plots established at a 2.44 m square spacing with a five-row external surround. Triple lattice designs were imposed at the principal localities to contend with variation that was expected to stem from the large number of families and the heterogeneity of experimental sites. From this trial, an average increase of 30 % in relative efficiency of the triple lattice design over RCB was reported (BARNES &

SCHWEPPENHAUSER 1979, BARNES *et al.* 1992a,b). In his analysis of a *Pinus banksianan* Lamb. family test with a cubic lattice design in Manitoba, Canada, Dr. Jerome Klein showed a reduction of the proportion of the plot error variance for 10-years height from 13.9 % (when analyzed as a randomized block experiment) to 3.1 % and an 11.6 % increase in individual heritability estimate (KLEIN 1989).

To assess the effectiveness of other incomplete block designs, we examined the relative statistical efficiencies of ICBs over their corresponding RCBs in estimating family means with a computer simulation of a half-sib progeny trial with single-tree plots on one test site with various environmental variations (FU *et al.* 1998). The ICBs with Full Constraint (FC), Half Constraint (HC), No Constraint (NC) over replicates, and an alpha design (see WHITAKER *et al.* 2002) were considered. A geostatistical spatial model was employed, which allowed specifications of patchy and systematic environmental variation. The simulation showed that the alpha design was the most efficient ICB in the 21 scenarios of patchy and systematic site variation, followed by NC and HC, and these ICBs (except FC) were generally more efficient than RCB in terms of the average variance of a family mean contrast. The gains in statistical efficiencies with the alpha design and NC (rather than RCB) can be high, depending largely on the nature (and level) of spatial environmental variation.

Such gain can be empirically illustrated with the B.C. Douglas-fir trials. If the alpha design had been originally applied to the trials, a statistical efficiency of 1.25 in estimating family means (relative to RCBs) could be derived from Table 3 of FU *et al.* 1998 when the patch size of 18 m across (FU *et al.* 1999b) was considered alone. With presence of both patch variations and gradients on the test sites, the relative efficiency could have been higher than 1.25.

Heterogeneous nature of site variations

Effectiveness of ICB in removing site variation will increase on test sites of highly heterogeneous environments. It has been long known (e.g., See BATCHELOR & REED 1918) that site variation is the norm in forest genetic field trials, as the areas used are usually quite large in size (2–4 hectares or more) and are often on slopes or terrain where environmental gradients (e.g., soil depth, drainage, etc.) and patchy microsite patterns in forest soils exist. Even on seemingly homogeneous sites, nutrient and water gradients can be substantial. However, less known is the nature (and level) of site variation (MAGNUSSEN 1990).

We conducted a study to examine spatial variation patterns of tree heights in a series of B.C. Douglas-fir progeny trials with conventional statistics and geostatistical techniques (FU *et al.* 1999b). We found that there were large variations in tree height over the years within and among the 66 test sites. The estimated proportions of the within-site variance explained by family, row, column, patchiness, and within-plot, were on average 11, 7, 5, 12, and 47 %, respectively, plus 7 % due to unknown factors (Table 1 of FU *et al.* 1999b). The detailed distributions of such proportions explained by row, column, patchiness and within-plot, in relation to the within-site variances at two ages over the 66 test sites, are shown in Fig. 2. Significant gradients in row and column directions were observed in more than 44 test sites and the estimated slopes ranged in average from 0.33 to 1.52 cm/plot. Patch sizes varied greatly over the test sites and ranged in average from 5.21 to 6.47 plots, indicating that the average patch size for these trials was 18 m across. Temporal variations were large for family variance, but not much for those variance proportions explained by row, column, patchiness and within-plot. When trees grew older, more significant gradients were found and larger patch sizes observed.

Similar patterns of site variation were also observed on the Ontario farm-field test of black spruce progeny (JOYCE *et al.* 2002). 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. 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.

These findings illustrate that the environmental variations on the test sites examined were large and exhibited complex patterns. These variations may be difficult to be modeled with simple gradients and spherical correlation. Clearly, such complexities provide another justification of the need to develop and implement smaller blocking that can better partition and account for more environmental variation in the progeny trials of forest trees.

Advantages of small blocking

In most field tests, it is desirable to choose trial sites that reflect the environments in which their progeny will be planted (NAMKOONG *et al.* 1988). Thus the patterns and magnitudes of site variation as observed in the Douglas-fir trials are largely expected, although they may differ quantitatively for different trials and test sites elsewhere. Considering the heterogeneous

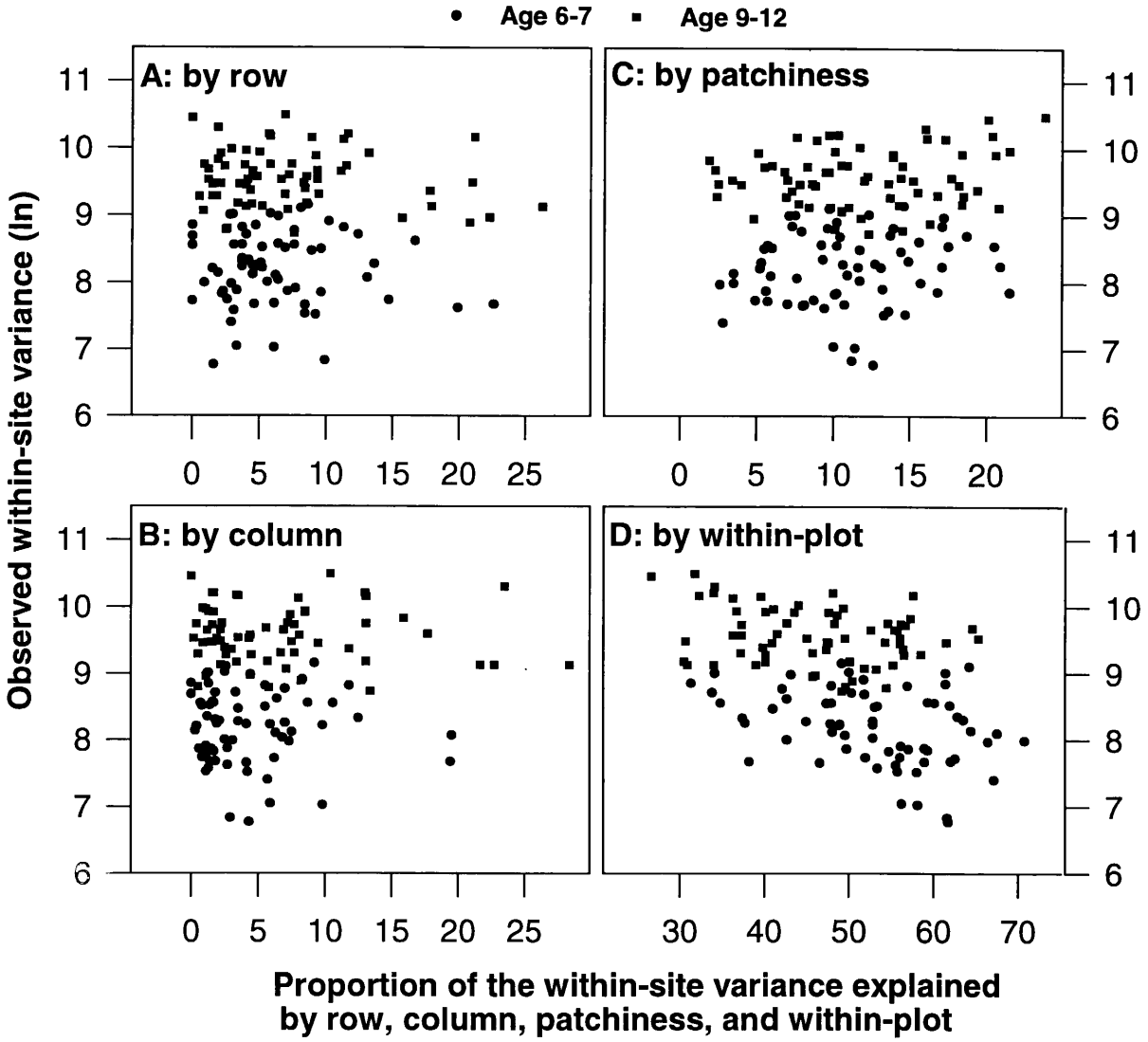

Figure 2. The distributions of the proportions of within-site variance explained by row, column, patchiness, and within-plot, in relation to the within-site variances of tree height observed twice (at age 6–7 and age 9–12) on the 66 test sites in a series of the Douglas-fir progeny trials (see FU *et al.* 1999b).

nature of site variation, there seems to be no doubt about the effectiveness of small blocking in reducing the effects of environmental heterogeneity and increasing the statistical efficiencies of estimating genetic values of breeding materials. This clearly reflects the major advantage small blocking has over RCBs. However, how much is gained will depend on many factors such as the type of blocking, the nature of spatial variation, and mortality rate.

In practice, small blocking allows flexibility in controlling site variation and choosing planting sites. With small blocks, tree breeders can even conduct genetic trials on irregular surface areas. This means that we can test on the higher elevation and rougher terrain that are more critical to establish for much testing. It then becomes feasible to obtain more precise estimates of key environmental effects since more extreme ranges can be sampled. Such gains in precision of estimates per site would enhance the studies on the nature of genotype-environment interactions reflected over many test sites (GREGORIUS & NAMKOONG 1986, WHITE 1996).

A gain in statistical efficiency means an opportunity to lower the experimental cost. A relative efficiency of 1.10 for ICB over RCB means that the same precision of estimating family means as for a standard design such as RCB can be obtained with ICB, but with 10 % fewer trees. Larger tree improvement programs start to save most in absolute terms, because testing is the most costly phase of tree breeding. With small blocking, a

direct cost reduction could also be appreciated, first in producing the trees, applying treatments or making crosses, then in planting, protecting, and measuring them. Thus, the combined economic benefits from implementing of ICBs should be far from trivial. A specific example for such economic benefit can be given with the Douglas-fir trials. We shown with computer simulation that ICBs, if applied for the Douglas-fir trials, would have achieved an efficiency of 1.25 or more (relative to the RCBs used) in estimating family means (FU *et al.* 1998, FU *et al.* 1999b). Thus, if considering family selection only, 25 % fewer trees could be used in the whole trials with ICBs for the same level of efficiency achieved currently with the RCBs used. This translates into a reduction of 52,800 trees (88 sites × 150 families × 16 trees per family × 25 %) and a saving of $528,000 (with a cost estimate of ~$10 per tree in the genetic testing).

ISSUES AND CONCERNS OF IMPLEMENTING SMALL BLOCKING

With small blocking, efforts will increase in field layout of testing materials and in data collection. Also, more efforts may be needed to analyse the data, as the data may be less balanced due to mortality and the analysis may require more advanced computer packages. More interestingly *and wrongly*, postblocking of completely randomized designs has been recommended for some practical reasons, rather than using efficient block designs in the first instance (ERICSSON 1997, WILLIAMS & FU 1999). There is no doubt that there are more issues and concerns than those mentioned. Thus it is worth discussing some of them here, at least from the practical point of view.

Efficiencies with different estimators and testing materials

One of the concerns expressed to us involved the statistical efficiencies with different estimators and testing materials, as our investigation applied mainly the estimates of family means to calculate relative efficiency for comparison of block designs and did not consider all the testing materials. In practice, however, estimates of family mean, heritability, or breeding value to evaluate family and/or individual performances are often made (NAMKOONG *et al.* 1988). When different estimators are used, the relative efficiency with small blocking over RCBs can significantly differ. Generally, estimates of family means are more sensitive to blocking methods than those of heritability and breeding values, which are functions of genetic variance (FU *et al.* 1998). However, there are no particular reasons why small blocking should be less efficient with different estimators than RCBs used on heterogeneous test sites (See SCHUTZ & COCKERHAM 1966 for discussion). For example, when the accuracy of family ranking (again, a function of heritability) was used to compare block designs, ICBs still showed a slight superiority over RCBs (FU *et al.* 2000).

There are several types of testing materials commonly used in genetic field trials (e.g., half-sib progeny, full-sib progeny, selfed progeny, clonal materials). For a given testing material (e.g., full-sib progeny), small blocking should outperform randomized complete block designs as long as the test sites are heterogeneous with gradients and patchy variations. However, the magnitude of increased efficiency of estimating family means might differ for various testing materials, because the within-family genetic variation differs among progeny for different testing materials and their confounding effects with various design parameters can increase and/or reduce the efficiency. In our investigation, we made comparisons in efficiency among half-sib, full-sib and clonal materials. We found that there were small differences in precision of mean estimations between full-sib family and clonal tests with respect to the design parameters examined (FU *et al.* 1999c), but some difference existed with presence of mortality in the trials (FU *et al.* 1999a).

Choice of incomplete block designs

There are many incomplete block designs tree breeders could apply (COCHRAN & COX 1957, HINKELMANN & KEMPTHORNE 1994, JOHN & WILLIAMS 1995, WILLIAMS *et al.* 1999). The choice of small blocking depends largely on the knowledge of site variation. If there are only one-dimensional gradients on most of the test sites, one may need to consider those designs with just one-way blocking, such as the alpha designs. If the site variation is more heterogeneous with two or more directional variations, it would be better to apply row and column designs. From the spatial analysis of the Douglas-fir progeny trials, it is clear that there were 2/3 of the test sites displaying significant gradients in row and column directions and that the average patch size for these trials was 18 meters across. Thus, row and column designs such as the latinized row and column designs (e.g., see JOHN & WILLIAMS 1998) should have been preferred in the Douglas-fir progeny trials.

In our investigation, we examined four randomized incomplete block designs, all of which are one-way blocking designs (i.e., intended to remove site variation in one direction). As shown in FU *et al.* (1998), the

alpha design was the best, followed by NC, HC, and FC, in controlling the site variations that were presumably known. The alpha design is a class of generalized lattice designs with more flexibility, i.e., they are available whenever the number of families is a multiple of the block size and they can be easily adapted even when it is not (JOHN & WILLIAMS 1995). Thus, they are generally suitable for forest genetic trials with a large number of families to be evaluated and provide better control of site variations. However, in some situations such as having unequal block or family sizes where the alpha design cannot be generated, alternatives are also available. Our simulations indicate that the randomized incomplete block designs with no restrictions on experimentation over replicates (i.e., NC) can be nearly as efficient as the alpha design.

Proper blocking on test sites

The commonly applied approach for blocking on a test site is first to make the best educated guess (or projection) of variation patterns on the site in terms of patch variations and gradients, and then based on this projection, to determine optimum design parameters and perform proper field layouts. For example, if large patch sizes are expected, block sizes may be adjusted accordingly. Realistically, however, the best projections may not always be achieved and the design parameters used may not necessarily be optimal. While projections largely depend on the experience a tree breeder has on test sites, the choice of design parameters and the field layout could be made best only with sufficient knowledge on the relations between the effectiveness of small blocking and site variation.

The choice of block size, shape and orientation, and family size has been studied (WRIGHT & FREELAND 1960, CONKLE 1963, JOHNSTONE & SAMUEL 1974, LIBBY & COCKERHAM 1980, LEE 1983, LAMBETH *et al.* 1983, COTTERILL & JAMES 1984, CORRELL & CELLIER 1987, LOO-DINKINS & TAUER 1987, MCCUTCHAN *et al.* 1989, HAAPANEN 1992), but most studies disregarded the variable nature of site variation. Even with considerations of site variation, these studies were usually site-specific (e.g., LAMBETH *et al.* 1983, CORRELL & CELLIER 1987, LOO-DINKINS & TAUER 1987). Also, most of these studies considered only RCBs, in which 4–6 blocks of larger than 150 individual trees each were usually used (MAGNUSSEN 1993a). With such large blocks arranged in squares or rectangles over test sites, one would expect, for example, that orientation of blocking might not be important in removing site variation. Such expectation may not always hold when blocks of size less than 20 are used.

While the information derived from large blocking is still useful, it would be more constructive to examine the effectiveness of various blocking of small sizes. We conducted several computer simulations of full-sib and clonal trials with single-tree plots to address some issues such as the appropriate block size, the imbalance in block size, and the effectiveness of row, column, or square blocks (FU *et al.* 1999c). These simulations showed that the rule of thumb for optimising the block size (i.e., approximately the square root of family number to be tested) should still apply (PATTERSON & HUNTER 1983). Slight variation in block size did not seem to have a large impact on the precision of estimating family means when blocks were considered random. Without gradients, column and row blocks were equally effective in removing patchy variations, but column blocks were more effective than row blocks in removing gradients in row direction. Square blocks were more effective than either column or row blocks in removing both patchy variations and gradients defined in one direction, but this still depends largely on the block sizes used. Thus, blocking in small squares should always be preferred, as the nature of site variation is rarely known with certainty. This reflects the notion of maximal compactness in field layout that was pointed out by Dr. Sally John.

Analysis of data from small blocking

Statistical analyses of data from incomplete block designs with and without the recovery of interblock information are extensively discussed in the literature (e.g., see PATTERSON & THOMPSON 1971, GIESBRECHT 1986, and WILLIAMS & MATHESON 1994). In general, the analyses without the recovery of interblock information (i.e., block effects are fixed) can cause many difficulties in complex designs and in the case of missing data. When block effects are considered random, these problems can be handled with the modified maximum likelihood principle proposed by Patterson and Thompson (1971), i.e., what is usually called the REML method. This seems to be true from our experience with single-site univariate ICB data (even with missing values). For estimation of family means, for example, analyses of these data with SAS® PROC MIXED (SAS INSTITUTE INC 1996) are no longer a barrier (FU *et al.* 1998, 1999a). SAS® PROC MIXED allows fits of mixed linear models (models with both fixed and random effects) and provides easy access to a variety of mixed models useful in many common statistical analyses, including split-plot designs, random coefficients, best linear unbiased prediction and heterogeneous variances. However, SAS®

PROC MIXED can not iterate to a solution for problems where a G matrix is specified and is also unable to iteratively solve multivariate genetic problems. This situation applies to multi-site data where estimators may be biased, especially when the genotype-environment interactions are strong (LU *et al.* 1999). This requires further research, but intuitively we have no special concerns over estimates from multiple sites, since in most cases estimates obtained from RCB designs would carry the same, or higher, biases.

For multi-site ICB data, one could still follow what White and Hodge (1992) have proposed (i.e., BLP-Best Linear Prediction or BLUP- Best Linear Unbiased Prediction), depending on specific testing situations. Whether the estimates obtained are BLP or BLUP depends largely on the use of models that reflect the applied block designs to estimate the fixed effects. If the data are analyzed first for fixed effects and then prediction of random genetic effects is made, the estimates of breeding value are probably BLP. When the estimates of fixed effects are simultaneously made with the REML method that takes into account the genotype-environment interactions, one would get BLUP. However, when the data set is large like those in forest trials, the REML method may not always be easy to implement in the analysis. This may require specific analysis packages such as SAS® PROC MIXED or GENSTAT programs in which REML routines are provided (SEARLE *et al.* 1992, GILMOUR *et al.* 1995).

Spatial analysis of data from either RCBs or small blocking represents another challenge (MAGNUSSEN 1990, 1993B, ANEKONDA & LIBBY 1996, JOYCE *et al.* 2002). This requires more understanding of environmental variation and more research on the nature (and level) of site variation (MAGNUSSEN 1990). In our investigation, we examined various patterns of site variation, first by analyzing the large-scale deterministic structures with median-polishing methods and then the small-scale stochastic structures with variography (CRESSIE 1991). Such two-steps analyses generated detailed information on both gradients and patchiness for genetic testing and showed advantages over other spatial analyses that do not separate trends. These analyses showed that the spherical covariance model seemed to fit well with the observed site variation, but this did not exclude the presence of other spatial models (see FU *et al.* 1999b). SAS® PROC VARIOGRAM and PROC NLIN should facilitate the assessments of various spatial variation models. However, how to incorporate these covariance models into data analyses remains to be examined, even though some research has been done for variety trials (e.g., see BESAG & KEMPTON 1986, ZIMMERMAN & HARVILLE 1991, CLARKE & BAKER 1996, CULLIS *et al.* 1998, APIOLAZA *et al.* 2000). Whether such spatial analysis will significantly enhance efficiency beyond that of incomplete block designs is of great interest and remains to be determined.

Computer programs

Dr. Emlyn Williams and his colleagues have actively developed many efficient designs and recently written various computer programs to generate various layouts of test materials (see WHITAKER *et al.* 2002; or http://www.ffp.csiro.au/software). Among those are the alpha designs and t-latinized designs (row and column designs). These programs generate not only the layouts, but also give the appropriate analysis models to be used (including analysis programs for SAS or GENSTAT). All of these will definitely facilitate the implementing of small blocking.

There are many computer packages now available for analysis of trial data, although not necessarily exclusively for data from small blocking. They include those procedures from SAS and GENSTAT programs, in which REML methods can be relatively easy to be implemented. SAS PROC MIXED and PROC LATTICE, in particular, are very useful in analysis of data from small blocking. Also, there are some PC programs available from animal breeders in which REML methods are incorporated into various BLUPs (e.g., DFREML by K. Meyer, MTDFREML by K. Boldman and D. van Vleck, ASReml by A. Gilmour). However, most of these programs are limited by the size of forest trail data (WHITE & HODGE 1988), while the ASReml algorithm seems to be powerful (APIOLAZA *et al.* 2000).

FUTURE DIRECTIONS

Genetic field trials are a vital part of forest genetic research and will continue to be critical for advances in tree improvement programs, particularly in advanced generations of tree breeding (ZOBEL & TALBERT 1984, NAMKOONG *et al.* 1988, MAGNUSSEN 1993a). Implementation of small blocking represents another attempt toward improving the informativeness of the future genetic trials, by reducing environmental variation to achieve higher accuracy and precision of estimating breeding values. Such implementation is supported from our investigation and related studies on small blocking and also shows promise of potential economic benefits, particularly with the advances in computer programs to generate efficient field layouts and analyse resulting data. To optimize such implementation,

further studies are still needed. Among those is extending our investigation to two or more factor multi-site experiments that can reflect the true picture of large genetic field trials. This may require the accurate modelling of genotype-environment interactions (e.g. see QIAO *et al.* 2000). Multi-factor experiments with multiple sites will generate more complex data sets that require more research on data analysis. Another extension is the consideration of partial diallel crosses in incomplete block designs which could be more effective (and informative) than the implementing of small blocking alone (NAMKOONG & ROBERDS 1974, BURDON & VAN BUIJTENEN 1990, SINGH & HINKELMANN 1995). Incorporation of spatial variation models in data analysis is another area of research that needs more attention, and could in turn improve genetic estimation (FU *et al.* 1999b, JOYCE *et al.* 2002). More analyses should be made of existing progeny trials to understand how the environmental variation patterns on test sites behave over time. This could help in the development of both efficient field designs and effective analytical methods.

ACKNOWLEDGMENTS

This paper was orally presented in the symposium of Unifying Perspectives of Evolution, Conservation and Breeding held in honour of Dr. Gene Namkoong at the University of British Columbia from July 22 to 24, 1999. It greatly benefited from the contributions made by Drs Gene Namkoong, Alvin Yanchuk, Peter Clarke and Emlyn Williams to many parts of my investigation on small blocking. It also benefited from the stimulating discussions, constructive suggestions and assistance in the investigation from Drs. Rowland Burdon, Klaus Hinkelmann, Sally John, John King, Mathew Koshy, Steen Magnussen, John Russell, William Wingle, Chang-Yi Xie, and Mr. Jack Woods.

REFERENCES

ANEKONDA, T. S. & LIBBY, W. J. 1996: Effectiveness of nearest-neighbor data adjustment in a clonal test of redwood. *Silvae Genetica* **45**: 46-51.

APIOLAZA, L.A., GILMOUR, A.R. & GARRICK, D.J. 2000: Variance modeling of longitudinal height data from a Pinus radiate progeny test. *Canadian Journal of Forest Research* **30**: 645-654.

BARNES, R.D. & SCHWEPPENHAUSER, M.A. 1979: *Pinus patula* Schiede and Deppe progeny tests in Zimbabwe Rhodesia: Genetic control of 1.5 year-old traits and a comparison of progeny test methods. *Silvae Genetica* **28**: 156-167.

BARNES, R. D., MULLIN, L. J. & BATTLE, G. 1992a: Genetic control in fifth year traits in *Pinus Patula* Schiede and Deppe. *Silvae Genetica* **41**: 242-248.

BARNES, R. D., MULLIN, L. J. & BATTLE, G. 1992b: Genetic control in eighth year traits in *Pinus patula* Schiede and Deppe. *Silvae Genetica* **41**: 318-326.

BATCHELOR, L. D. & REED, H. S. 1918: Relation of the variability of yields of fruit trees to the accuracy of field trials. *Journal of Agricultural Research* **12**: 245-285.

BESAG, K. & KEMPTON, R. 1986: Statistical analysis of field experiments using neighboring plots. *Biometrics* **42**: 231-251.

BURDON, R. D. & VAN BUIJTENEN, J. P. 1990: Expected efficiencies of mating designs for reselection of parents. *Canadian Journal of Forest Research* **20**: 1664-1671.

BURLEY, J., BURROWS, P.M., ARMITAGE, F.B. & BARNES, R.D. 1966: Progeny test designs for *Pinus patula* in Rhodesia. *Silvae Genetica* **15**: 166-173.

CLARKE, F.R. & BAKER, R.J. 1996: Spatial analysis improves precision of seed lot comparisons. *Crop Science* **36**: 1180-1184.

CLARKE, G. P., HAINES, L. M. & FU, Y. B. 1997: Kriging and field experiments. *In*: Proceedings of 51st session of the International Statistical Institute, book 2, International Statistical Institute, Istanbul, pp. 229-232.

COCHRAN, W. G. & COX, G. M. 1957: Experimental Designs, 2nd edition, Wiley, New York, 611 pp.

CONKLE, M. T. 1963: The determination of experimental plot size and shape in loblolly and slash pines. North Carolina State University, School of Forestry Technology Report No. 17.

CORRELL, R. L. & CELLIER, K. M. 1987: Effects of plot size, block size and buffer rows on the precision of forest trials. *Australian Forest Research* **17**: 11-18.

COTTERILL, P. P. & JAMES, J. W. 1984: Number of offspring and plot sizes required for progeny test. *Silvae Genetica* **33**: 203-209.

CRESSIE, N. A. C. 1991: Statistics for Spatial Data. John Wiley & Sons, Inc. New York, 900 pp.

CULLIS, B., GOGEL, B., VERBYLA, A. & THOMPSON, R. 1998: Spatial analysis of multi-environment early generation variety trials. *Biometrics* **54**: 1-8.

ERICSSON, T. 1997: Enhanced heritabilities and best linear unbiased predictors through appropriate blocking of progeny trials. *Canadian Journal of Forest Research* **27**: 2097-2101.

FRIEDMAN, S. & NAMKOONG, G. 1987: Estimating family means using unbalanced incomplete blocks. *In*: Proceedings of IUFRO Joint Meeting of Working Parties on Breeding Theory, Progeny Testing, and Seed Orchards, Williamsburg, USA (1986). North Carolina State University, Industrial Cooperative Tree Improvement Program, Raleigh, N.C. pp. 457-468.

FU, Y.B., CLARKE, G.P.Y., NAMKOONG, G. & YANCHUK, A.D. 1998: Incomplete block designs for genetic testing: statistical efficiencies of estimating family means. *Canadian Journal of Forest Research* **28**: 977-986.

FU, Y. B., YANCHUK, A. D., NAMKOONG, G. & CLARKE, G. P. Y. 1999a: Incomplete block designs for genetic testing: statistical efficiencies with missing observations. *Forest Science* **45**: 374-380.

FU, Y. B., YANCHUK, A. D. & NAMKOONG, G. 1999b: Spatial patterns of tree height variations in a series of Douglas-fir progeny trials: implications for genetic testing. *Canadian*

Journal of Forest Research **29**: 714 -723.

FU, Y. B., YANCHUK, A. D. & NAMKOONG, G. 1999c: Incomplete block designs for genetic testing: some practical considerations. *Canadian Journal of Forest Research* **29**: 1871-1878.

FU, Y. B., YANCHUK, A. D. & NAMKOONG, G. 2000: Incomplete block designs for genetic testing: accuracy of ranking families and individuals. *Forest Genetics* **7**: 287 -293.

GIESBRECHT, F.G. 1986: Analysis of data from incomplete - block designs. *Biometrics* **42**: 437-448.

GILMOUR, A. R., THOMPSON, R. & CULLIS, B. R. 1995: AI, an efficient algorithm for REML estimation in linear mixed models. *Biometrics* **51**: 1440 -1450.

GREGORIUS, H.-R. & NAMKOONG, G. 1986: Joint analysis of genotypic and environmental effects. *Theoretical and Applied Genetics* **72**: 413-422.

HAAPANEN, M. 1992: Effect of plot size and shape on the efficiency of progeny tests. *Silva Fennica* **26**: 201-209.

HEAMAN, J. C. 1978: Choosing strategies for a breeding program in Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco] from British Columbia. *In*: Proceedings of the 3rd world consultation on forest tree breeding. Canberra. Australia. pp. 1205-1214.

HINKELMANN, K. & KEMPTHORNE, O. 1994: Design and Analysis of Experiments Vol. 1. Introduction to Experimental Design. John Wiley & Sons, New York, 512 pp.

JOHN, J. A. & WILLIAMS, E. R. 1995: Cyclic and computer generated designs. Chapman and Hall, London, 256 pp.

JOHN, J. A. & WILLIAMS, E. R. 1998: *t*-latinized designs. *Australian & New Zealand Journal of Statistics* **40**: 111-118.

JOHNSTONE, R. C. B. & SAMUEL, C. J. A. 1974: Experimental design for forest tree progeny tests with particular reference to plot size and shape. *In*: Proceedings of the joint IUFRO meeting, S.02.04.1-3, Stockholm, Session IV, pp. 357-376.

JOYCE, D., FORD, R. & FU, Y. B. 2002: Spatial patterns of tree height variations in a black spruce farm-field progeny test and neighbors-adjusted estimations of genetic parameters. *Silvae Genetica* **51**: 13-18.

KLEIN, J. I. 1989: Effects of incomplete block design and stand structure on heritability of growth in a jack pine family test in Manitoba. *Canadian Journal of Forest Research* **19**: 1044-1050.

LAMBETH, C. C., GLADSTONE, W. T. & STONECYPHER, R. W. 1983: Statistical efficiency of row and noncontiguous family plots in genetic tests of loblolly pine. *Silvae Genetica* **32**: 24-28.

LEE, C. H. 1983: Statistical efficiency varies with plot size, number of replications and seedlots sampled. *Silvae Genetica* **32**: 49-52.

LIBBY, W. J. & COCKERHAM, C. C. 1980: Random noncontiguous plots in interlocking field layouts. *Silvae Genetica* **29**: 5-6.

LOO-DINKINS, J. 1992: Field test design. *In*: Handbook of quantitative forest genetics. (ed. L. Fins, S. Friedman & J.V. Brotschol). Kluwer Acadamic Publisher, Boston, pp. 96-139.

LOO-DINKINS, J. A. & TAUER, C. G. 1987: Statistical efficiency of six progeny test field designs on three loblolly pine (*Pinus taeda* L.) site types. *Canadian Journal of Forest Research* **17**: 1066 -1070.

LU, P. X., HUBER, D. A. & WHITE, T. L. 1999: Potential biases of incomplete linear models in heritability estimation and breeding value prediction. *Canadian Journal of Forest Research* **29**: 724-736.

MAGNUSSEN, S. 1990: Application and comparison of spatial models in analyzing tree-genetics field trials. *Canadian Journal of Forest Research* **20**: 536-546.

MAGNUSSEN, S. 1993a: Design and analysis of tree genetic trials. *Canadian Journal of Forest Research* **23**: 1144 -1149.

MAGNUSSEN, S. 1993b: Bias in genetic variance estimates due to spatial autocorrelation. *Theoretical and Applied Genetics* **86**: 349-355.

MCCUTCHAN, B. G., NAMKOONG, G. & GIESBRECHT, F. G. 1989: Design efficiencies with planned and unplanned unbalance for estimating heritability in forestry. *Forest Science* **35**: 801-815.

MCCUTCHAN, B. G., OU, J. X. & NAMKOONG, G. 1985: A comparison of planned unbalanced designs for estimating heritability in perennial crops. *Theoretical and Applied Genetics* **71**: 536-544.

NAMKOONG, G. & ROBERDS, J. 1974: Choosing mating designs to efficiently estimate genetic variance components for trees. *Silvae Genetica* **23**: 43-54.

NAMKOONG, G., KANG, H. C. & BROUARD, J. S. 1988: Tree Breeding Principles and Strategies. Monographs on Theoretical and Applied Genetics. Springer-Verlag, New York, 180 pp.

PATTERSON, H. D. & HUNTER, E. A. 1983: The efficiency of incomplete block designs in National List and Recommended List cereal variety trials. *Journal of Agricultural Science* **101**: 427-433.

PATTERSON, H. D. & THOMPSON, R. 1971: Recovery of interblock information when block sizes are unequal. *Biometrika* **58**: 545-554.

QIAO, C. G., BASFORD, K. E., DELAY, I. H. & COOPER, M. 2000: Evaluation of experimental designs and spatial analysis in wheat breeding trials. *Theoretical and Applied Genetics* **100**: 9-16.

SAS® INSTITUTE INC. 1996: SAS® User Guide. Version 6.03 Ed. Cary, NC, 1028 pp.

SCHUTZ, W. M. & COCKERHAM, C. C. 1966: The effect of field blocking on gain from selection. *Biometrics* **22**: 843-862.

SEARLE, S. R., CASELLA, G. & MCCULLOCH, C. E. 1992: Variance Components. John Wiley & Sons, Inc. New York, 501 pp.

SINGH, M. & HINKELMANN, K. 1995: Partial diallel crosses in incomplete block designs. *Biometrics* **51**: 1302-1314.

WHITAKER, D., WILLIAMS, E. R. & JOHN, J. A. 2002. CycDesigN: A package for the computer generation of experimental designs. CSIRO, Canberra.

WHITE, T. 1996: Genetic parameter estimates and breeding value prediction: issues and implications in tree improvement programs. *In:* Proceedings of QFRI-IUFRO Conference on Tree Improvement For Sustainable Tropic Forestry (ed. M.J. Dieters, A.C. Matheson, D.G. Nikles, C.E. Hardwood & S.M. Walker), Caloundra, Queensland, Australia. pp.110-117.

WHITE, T. L. & HODGE, G. R. 1988: Best linear prediction of

breeding values in a forest tree improvement program. *Theoretical and Applied Genetics* **76**: 719-727.

WHITE, T. L. & HODGE, G. R. 1992: Predicting Breeding Values with Applications in Forest Tree Improvement (Forestry Sciences Series vol. 33). Kluwer Academic Publishers, Dordrecht, 363 pp.

WILLIAMS, E. R. & MATHESON, A. C. 1994: Experimental Design and Analysis for Use in Tree Improvement. CSIRO Publication, Australia, 174 pp.

WILLIAMS, E. R. & FU, Y. B. 1999: Comment on enhanced heritabilities and best linear unbiased predictors through appropriate blocking of progeny trials. *Canadian Journal of Forest Research* **29**: 1633-1634.

WILLIAMS, E. R., JOHN, J. A. & WHITAKER, D. 1999: Example of block designs for plant and tree breeding trials. *Australian & New Zealand Journal of Statistics* **41**: 277-284.

WRIGHT, J. W. & FREELAND, F. D. 1960: Plot size and experimental efficiency in forest genetic research. Mich. State Univ. Agric. Exp. Stn. Tech. Bull. No. 280, 28 pp.

YANCHUK, A. D. 1996: General and specific combining ability from disconnected partial diallels of coastal Douglas-fir. *Silvae Genetica* **45**: 37-45.

YATES, F. 1936: A new method of arranging variety trials involving a large number of varieties. *Journal of Agricultural Science* **26**: 424-455.

ZIMMERMAN, D. L. & HARVILLE, D. A. 1991: A random field approach to the analysis of field-plot experiments and other spatial experiments. *Biometrics* **47**: 223-239.

ZOBEL, B. & TALBERT, J. 1984: Applied Forest Tree Improvement. John Wiley & Sons, New York, 505 pp.